

In dit vierde deel van de serie *AI in de Adviespraktijk* richten we ons op een cruciaal aspect: hoe weet en meet je of je AI-systeem daadwerkelijk doet wat het moet doen? In het eerste deel (VVP 1, 2024) liet AI-strateeg Dennie van den Biggelaar zien hoe je van start gaat met Machine Learning (een specifiek onderdeel van AI), in het tweede deel (VVP 2) hoe je AI operationaliseert in je bedrijfsprocessen en in het derde deel (VVP 3) stond Integreren van AI-software in bestaande IT-landschappen centraal.

## Meten = leren

TEKST DENNIE VAN DEN BIGGELAAR, ONESURANCE

Het meten van de effectiviteit van een AI-toepassing begint met het definiëren van duidelijke 'business KPI's'. Deze KPI's zijn essentieel omdat ze richting geven aan welke aspecten van je bedrijfsvoering je wilt verbeteren en hoe je deze verbeteringen meetbaar kunt maken. Voor een verzekeringsbedrijf kunnen deze doelen bijvoorbeeld zijn: verhogen van omzet, verbeteren van retentie, verhogen van polisdichtheid of verhogen van STP-acceptatie. Het vaststellen van deze KPI's biedt een kader voor zowel de ontwikkeling als de evaluatie van de AI-applicatie.

### MENS EN MACHINE

In de praktijk werken AI-toepassingen vaak samen met menselijke experts. Daarom is het belangrijk om de prestaties van zowel de AI als de mens afzonderlijk én gezamenlijk te meten. Dit geeft inzicht in de effectiviteit van de samenwerking en helpt je te bepalen waar verbeteringen mogelijk zijn.

**Voorbeeld: Actief klantbeheer:** stel je hebt een AI-algoritme dat klanten identificeert met een hoge kans op roeyement. Als de binnendienst of adviseur deze signalen niet adequaat opvolgt, kan de beoogde reductie van roeyement uitblijven. Door prestaties per medewerker te meten, kun je ontdekken of bepaalde medewerkers betere resultaten boeken dan anderen. Deze inzichten kunnen vervolgens worden gedeeld om het team als geheel te versterken.

### TECHNISCHE PERFORMANCE

Om de technische performance van een voorspellend algoritme te beoordelen, worden verschillende indicatoren gebruikt: nauwkeurigheid of accuracy (geeft aan hoe vaak het algoritme de juiste voorspelling doet), precisie (deze maatstaf kijkt specifiek naar de betrouwbaarheid van positieve voorspellingen), sensitiviteit (dit meet hoe goed een AI-model in staat is om alle relevante uitkomsten te detecteren), Area Under Curve (biedt een overzicht van de voorspellingskwaliteit van het model over verschillende

drempelwaarden) en Log Loss ofwel Logaritmisch Verlies (dit meet hoe dicht de voorspelde kansen bij de werkelijke uitkomsten liggen).

Naast deze indicatoren zijn snelheid, efficiëntie en schaalbaarheid van belang. De snelheid, oftewel latency, bepaalt hoe snel de AI-applicatie reageert op een verzoek. Efficiëntie wordt gemeten door het geheugenverbruik van de applicatie, en schaalbaarheid wordt beoordeeld op basis van de hoeveelheid voorspellingen die binnen een bepaalde tijd worden gerealiseerd (throughput). Deze factoren geven een oordeel over de schaalbaarheid van een algoritme.

### ROBUUST EN ETHISCH

Een AI-toepassing moet niet alleen technisch goed presteren, maar ook robuust en ethisch verantwoord zijn. Dit omvat het vermogen van het model om goed te blijven presteren, zelfs als de inputdata of de omgeving verandert (model drift en shift). Daarnaast moet het model gevoelig zijn voor veranderingen in de data

**'Meten effectiviteit van een AI-toepassing is een complex maar noodzakelijk proces'**

waarop het is getraind (data drift en shift). Ethische overwegingen, zoals het voorkomen van discriminatie op basis van geslacht, etniciteit of leeftijd, zijn eveneens cruciaal om te waarborgen dat de AI eerlijk en verantwoord opereert.

### UPTIME EN BETROUWBAARHEID

Zoals bij elke cloudgebaseerde applicatie is de uptime van een AI-toepassing van cruciaal belang, vooral in productieomgevingen. Een gebruikelijke norm in een Service Level Agreement (SLA) is een uptime van 99,9 procent. Dit betekent dat van elke 1.000 interacties met de applicatie er niet meer dan één fout mag gaan. Om deze betrouwbaarheid te waarborgen, wordt vaak een reserve-applicatie ingezet die het eventueel kan overnemen bij uitval.

### VAN PROTOTYPE NAAR PRODUCTIE

Het opzetten van een AI-toepassing is een stapsgewijs proces. In de prototypefase ligt de focus vooral op het testen van de voorspelbaarheid van het algoritme en het minimaliseren van eventuele discriminatie. Als de AI-applicatie deze tests doorstaat, is de volgende stap om te beoordelen of de toepassing daadwerkelijk de gewenste business KPI's verbetert. Ook de schaalbaarheid van het model wordt in deze fase al in overweging genomen.

Wanneer de AI eenmaal in productie is, verschuift de focus naar het waarborgen van de uptime en het monitoren van de robuustheid van de AI over de tijd. Door systematisch te meten en te evalueren, kun je continu verbeteren en ervoor zorgen dat je AI-toepassing doet wat zij moet doen, nu en in de toekomst.

### METEN VAN IMPACT

Een van de meest effectieve methoden om te meten of een AI-toepassing de gewenste resultaten oplevert, is door middel van A/B-testing. Hierbij wordt de doelgroep willekeurig verdeeld in twee groepen: de ene groep (Groep A) gebruikt de nieuwe AI-toepassing, terwijl de andere groep (Groep B) de traditionele methode of een eerdere versie van het systeem zonder AI gebruikt. Door de prestaties van beide groepen te vergelijken, kun je vaststellen hoe effectief de AI is in het verbeteren van de business KPI's.

Het succes van een AI-toepassing hangt sterk af van hoe de inzichten uit A/B-tests worden geïntegreerd in de bedrijfsvoering. Als een A/B-test bijvoorbeeld laat zien dat een bepaal-

*Dennie van den Biggelaar:  
'Effectief bijdagen aan verbeterde bedrijfsresultaten.'*



de AI-tool leidt tot een hogere polisdichtheid, kan dit aanleiding zijn om de tool breder uit te rollen binnen de organisatie.

### EFFECTIEF

Het meten van de effectiviteit van een AI-toepassing is een complex maar noodzakelijk proces. Het begint bij het definiëren van heldere business KPI's en het evalueren van zowel de technische prestaties als de samenwerking tussen mens en machine. Robuustheid, ethische overwegingen en uptime zijn net zo belangrijk als de voorspelbaarheid van het algoritme. Door gebruik te maken van A/B-testing kun je bovendien op een betrouwbare manier vaststellen of de AI-applicatie daadwerkelijk bijdraagt aan het behalen van je bedrijfsdoelen. Het is van essentieel belang dat het niet alleen technisch goed functioneert, maar ook effectief bijdraagt aan het verbeteren van je bedrijfsresultaten. ■

### AI IN DE ADVIESPRAKTIJK

In deze rubriek legt AI strateeg Dennie van den Biggelaar (Onesurance) uit hoe je specifieke AI en machine learning toepast op 'advies in de praktijk'. In verschillende edities van VVP worden de volgende onderwerpen uitgelicht:

- Starten met specifieke AI en ML
- Operationaliseren in bedrijfsprocessen
- Integreren in bestaand IT landschap
- Meten = leren: KPI's voor ML
- Ethiek, regelgeving en maatschappij
- AI en ML: een kijkje in de nabije toekomst