

Artificiële Intelligentie: technieken en toepassingen bij verzekeraars en DNB

Technologische innovatie is één van de belangrijkste ontwikkelingen in de financiële sector. Een belangrijk onderdeel hiervan is het toenemende gebruik van Artificiële Intelligentie (AI). Dit biedt kansen voor actuarissen, aangezien de (kwantitatieve) kennis en vaardigheden die nodig zijn voor het omgaan met AI, voor een groot deel overeen komen met de kennis en vaardigheden van een actuaaris.

In dit artikel geven wij een overzicht van wat AI inhoudt, staan wij stil bij hoe verzekeraars AI kunnen toepassen, en geven wij voorbeelden van hoe wij bij DNB voor het toezicht op verzekeraars gebruik maken van AI.



ACHTERGROND

De afgelopen jaren heeft DNB via verschillende onderzoeken aandacht besteed aan technologische innovatie in de Nederlandse verzekeringssector en de impact hiervan op de strategie, het bedrijfsmodel en de bedrijfsvoering van verzekeraars. Hieruit kwam naar voren dat technologische innovatie voor verzekeraars zowel kansen als risico's biedt. Gebieden waar voor verzekeraars kansen liggen zijn onder andere:

- Ontwikkeling van nieuwe producten en diensten
- Klantcontact en -behoud
- Kostenefficiëntie en schadelastreductie

Een belangrijke technologische ontwikkeling bij verzekeraars is het toenemende gebruik van AI.



S.R. Altemühl MAST AAG (links) is toezichthouder specialist bij het expertisecentrum Kapitaal en Actuarieel van de divisie Toezicht Verzekeraars van De Nederlandsche bank.

A.W.M. van Ool MSc is toezichthouder specialist bij het expertisecentrum Data-analyse en Bedrijfsmodellen van de divisie Toezicht Verzekeraars van De Nederlandsche bank.

Dit artikel is geschreven op persoonlijke titel.



OVERZICHT VAN AI

Artificiële intelligentie is intelligentie van computers. Het is een verzamelterm en verwijst naar een grote hoeveelheid technieken én toepassingen.

Helaas worden er binnen AI niet altijd eenduidige definities gebruikt en is er veel overlappend jargon. Dit komt omdat in het verleden vanuit verschillende vakgebieden onderzoek naar AI is gedaan (bijvoorbeeld vanuit engineering, computerwetenschappen en econometrie), en deze vakgebieden pas in recente jaren meer met elkaar in aanraking zijn gekomen.

In de context van AI wordt vaak ook gesproken over machine learning, waarmee doorgaans wordt verwezen naar de statistische algoritmes waar binnen AI gebruik van wordt gemaakt. Daarnaast worden patroonherkenning/data mining en natural language processing gezien als belangrijke toepassingsgebieden van AI. We lichten deze begrippen hieronder nader toe.

MACHINE LEARNING

Machine learning algoritmes herkennen verbanden in historische data, en gebruiken dit om een specifieke taak uit te voeren of een voorspelling te doen. In het bedrijfsleven wordt machine learning soms *predictive modelling* of *predictive analytics* genoemd.

Bij machine learning wordt vaak onderscheid gemaakt tussen supervised learning en unsupervised learning, waarbij de onderverdeling afhangt van het type data waarop een algoritme wordt getraind.

Supervised learning

Bij supervised learning wordt gebruik gemaakt van gelabelde data, en is het doel om op basis van een reeks *features* (onafhankelijke variabelen, x) te voorspellen wat de waarde is van de *label* (afhankelijke variabele, y).

De volgende technieken zijn voorbeelden van supervised machine learning:

- Lineaire regressiemodellen
- Beslisbomen
- Neurale netwerken (ook wel *deep learning*)¹

In de verzekeringssector wordt al meerdere decennia gebruik gemaakt van lineaire regressiemodellen. Het gebruik van beslisbomen en neurale netwerken is in opkomst, aangezien deze technieken beter in staat zijn om complexe patronen te herkennen in grote hoeveelheden data.

Unsupervised learning

Bij unsupervised learning is het doel om verbanden te herkennen in een dataset zonder labels (alleen x , geen y).

De volgende technieken zijn voorbeelden van unsupervised machine learning:

- clustering
- anomalie detectie

Clustering algoritmes verdelen observaties in een dataset in groepen met vergelijkbare eigenschappen. Een voorbeeld van clustering is het k -means algoritme, wat een dataset opsplijt in k vergelijkbare groepen.

Anomalie detectie algoritmes proberen niet om observaties in een dataset te groeperen, maar om individuele observaties te identificeren die materieel afwijken van de rest van de dataset (anomalieën/

uitschieters). Een voorbeeld van een anomalie detectie algoritme is *isolation forest*.

PATROONHERKENNING/DATA MINING

Een toepassingsgebied binnen AI is patroonherkenning/data mining, waarbij het doel is om in grote datasets terugkerende patronen te herkennen. De term patroonherkenning heeft zijn oorsprong in engineering, data mining in computerwetenschappen.

Bij patroonherkenning/data mining wordt gebruik gemaakt van breed inzetbare machine learning algoritmes zoals clustering en anomalie detectie, maar ook van specialistische *rule-based machine learning* technieken zoals *association rule learning*, die als doel hebben om in een dataset terugkerende regels² te herkennen.

NATURAL LANGUAGE PROCESSING

Een ander toepassingsgebied binnen AI is Natural Language Processing (NLP). Bij NLP is het doel om grote hoeveelheden natuurlijke taal (vaak tekstdocumenten) op geautomatiseerde wijze te verwerken en te analyseren, met als doel om hier bepaalde informatie uit te halen. Er bestaat een grote verscheidenheid aan NLP-technieken. Dit kan zowel supervised als unsupervised learning zijn.

Een voorbeeld van een supervised NLP-techniek is tekstclassificatie, waarbij labels aan tekstpassages worden toegekend. Op basis van een getrainde set van geclassificeerde voorbeelden wordt een machine learning model getraind, dat vervolgens wordt gebruikt om nieuwe tekstpassages te classificeren. Tekstclassificatie wordt onder andere gebruikt bij klantondersteuning, of voor het identificeren van spam e-mails.

Een voorbeeld van een populaire unsupervised NLP-techniek is topic modeling, waarbij tekst wordt gecategoriseerd naar onderwerpen die niet vooraf zijn gespecificeerd.

AI BIJ VERZEKERAARS

Nederlandse verzekeraars maken in toenemende mate gebruik van complexere AI modellen, vaak met als doel om bestaande (deel)processen te automatiseren of optimaliseren. Voorbeelden van toepassingen waarbij verzekeraars in toenemende mate gebruik maken van AI zijn fraudedetectie bij schadeclaims, het voorspellen van klantvragen, en het beter en sneller inschatten van schade-omvang. Voor de komende jaren wordt een sterkere focus verwacht op het toepassen van AI in andere gebieden, zoals premiestelling en acceptatie.³

Hieronder staan wij kort stil bij de toepassingen fraudedetectie en premiestelling.

Fraudedetectie

Een belangrijke AI toepassing voor verzekeraars is het beter opsporen van fraude bij schadeclaims. Hierbij kan gebruik worden gemaakt van anomaliedetectie om afwijkingen in claimgedrag te signaleren. Ook kan bij fraudedetectie gebruik worden gemaakt van netwerkalgoritmes waarmee verdachte relaties kunnen worden achterhaald.

Naast informatie over claimkenmerken maken verzekeraars bij fraude-detectie ook op grotere schaal gebruik van aanvullende databronnen, zoals foto's of prijslijsten van auto-onderdelen.

Premiestelling

Met behulp van nieuwe databronnen en krachtigere modellen kunnen verzekeraars op een gedetailleerder niveau risicogroepen onderscheiden, wat kan doorwerken in de premiestelling.

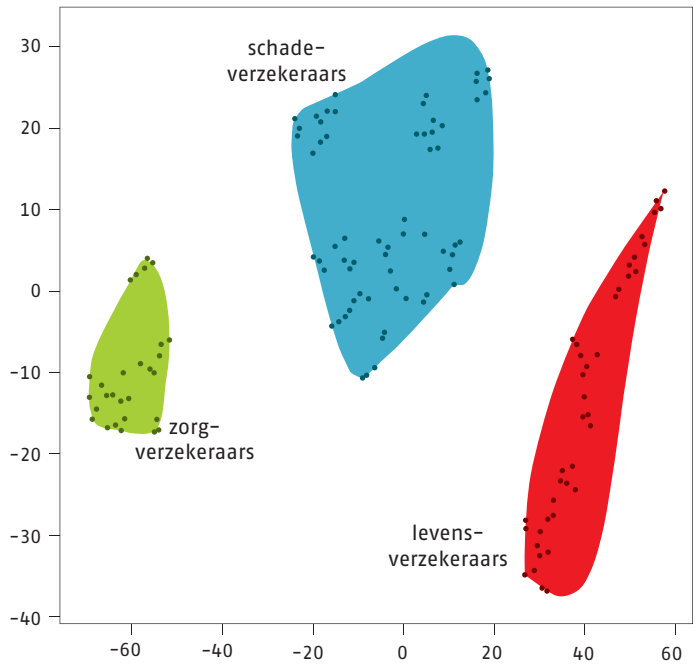
Op het gebied van het inschatten en beprijzen van risico's hebben verzekeraars de afgelopen jaren complexere AI modellen ingezet, al zij het beperkt. Voorbeelden hiervan zijn te vinden in het schadesegment (autoverzekeringen), en in mindere mate in het inkomenssegment. Op dit gebied worden in de komende jaren echter verdere stappen verwacht. Een mogelijkheid hierbij is dat verzekeraars voor het vaststellen van de 'actuariële' prijs nog gebruik blijven maken van traditionele modellen, zoals gegeneraliseerde lineaire regressie-modellen (GLM), maar voor het vaststellen van een commerciële prijs overstappen op complexere AI modellen.

AI BIJ DNB

Data is een cruciaal middel voor goed en efficiënt toezicht. Voor het analyseren van onze data maken wij bij DNB in toenemende mate gebruik van AI. Hieronder gaan wij kort in op een aantal voorbeelden van hoe wij bij DNB gebruik maken van AI. De broncode van deze voorbeelden is publiek beschikbaar op onze GitHub-pagina⁴. Verzekeraars kunnen deze broncode downloaden en toepassen op hun eigen data.

Clustering

Het k -means algoritme is een bekend voorbeeld van een unsupervised machine learning algoritme. Wij hebben het k -means algoritme toegepast op de publieke dataset met Solvency II gegevens van Nederlandse verzekeraars die DNB jaarlijks publiceert. Deze dataset bevat onder andere informatie over de balans, premies, technische voorziening en kapitaalvereisten van individuele verzekeraars. Door toepassing van het k -means algoritme worden verschillende clusters van verzekeraars geïdentificeerd, waarbij het cluster voornamelijk wordt bepaald door het type verzekeraar (zie onderstaande figuur).



Patroonherkenning en data mining

Het is belangrijk dat de data die wij ontvangen en gebruiken in ons toezicht van hoge kwaliteit is. Het controleren van de kwaliteit van grote datasets is echter een intensief proces.

Wij gebruiken patroonherkenning om het controleren van de datakwaliteit te versnellen. Met patroonherkenningalgoritmes kunnen op geautomatiseerde wijze controleregels voor datasets worden gegenereerd. Nadat de gegenereerde controleregels door een expert

zijn gereviewd, gebruiken wij deze om de rapportagedata die wij van verzekeraars ontvangen te controleren.

Het voordeel van patroonherkenningalgoritmes is dat controleregels niet handmatig hoeven worden opgesteld, maar automatisch worden gegenereerd. Hierdoor is het mogelijk om een grote hoeveelheid aan controleregels toe te passen gericht op het verbeteren van de datakwaliteit.

Recent hebben wij in een pilot met enkele verzekeraars en pensioenfondsen geëxperimenteerd met het toepassen van automatisch gegenereerde controleregels. In deze pilot gebruikten instellingen deze controleregels om hun eigen rapportagedata te controleren, voordat deze werd ingestuurd naar DNB (gebaseerd op het *first time right principe*).

Natural Language Processing

In ons toezicht hebben wij vaak ook te maken met tekstdata (bijvoorbeeld jaarverslagen, RSR's en ORSA's). Met behulp van NLP-technieken kunnen wij dergelijke documenten classificeren, of op geautomatiseerde wijze relevante informatie en passages uit deze documenten halen. Dit helpt onze toezichthouders bij het beoordelen van deze documenten.

AFSLUITING

Het toenemende gebruik van AI in de financiële sector is een belangrijke ontwikkeling, die voor verzekeraars zowel kansen als risico's biedt. Ook bij DNB zien wij kansen om bij het analyseren van onze data vaker gebruik te maken van AI. De komende jaren houden wij dan ook onverminderd aandacht voor de ontwikkelingen op dit gebied. ■

Bronnen

Insurtech – Een inventarisatie bij kleine en middelgrote verzekeraars (DNB, april 2020)

Artificiële Intelligentie in de verzekeringssector – een verkenning (AFM en DNB, juli 2019)

Visie op toezicht 2021–2024 (DNB, november 2020)

<https://github.com/DeNederlandscheBank>

Hulp bij ontrafelen fraude: netwerkanalyse (De Actuaris, februari 2020)

¹ – Alhoewel neurale netwerken vaak worden getraind op basis van gelabelde data (supervised learning), zijn er ook technieken voor het trainen van neurale netwerken op datasets zonder labels (unsupervised learning).

² – Dergelijke regels nemen typisch de vorm van een {ALS:DAN} uitdrukking, bijvoorbeeld {ALS 'rood' EN 'achthoekig' DAN 'stopbord'}.

³ – Zie voor meer informatie "Artificiële Intelligentie in de verzekeringssector – een verkenning" (AFM en DNB, juli 2019)

⁴ – <https://github.com/DeNederlandscheBank> (GitHub is een online platform waarop software geplaatst kan worden)