Artificiële Intelligentie in de verzekeringssector

een verkenning

DeNederlandscheBank
EUROSYSTEEM

AFM

1. Waarom deze verkenning?

2. Ontwikkelingen

3. Aandachtspunten

4. Effect van AI op solidariteit

Samenvatting

De ontwikkelingen omtrent Artificiële Intelligentie (AI) zijn in recente jaren wereldwijd in een stroomversnelling geraakt. Ook in de Nederlandse verzekeringssector wordt AI in toenemende mate ingezet. Deze ontwikkeling biedt legio kansen voor Nederlandse verzekeraars, maar brengt ook onzekerheden en risico's met zich mee. Voldoende aandacht voor en bewustzijn van die onzekerheden en risico's bij verzekeraars is cruciaal om te kunnen borgen dat AI ingezet wordt op een manier die verantwoord en in lijn is met de eisen omtrent integere en beheerste bedrijfsvoering, productontwikkeling en de zorgplicht.

Huidige AI-toepassingen

Momenteel experimenteren Nederlandse verzekeraars voorzichtig met zowel zelfontwikkelde en extern-ingekochte AI-toepassingen, veelal met het doel bestaande (deel)processen te automatiseren of optimaliseren. De gebruikte AI-toepassingen vallen onder 'narrow' AI: ze zijn gefocust op een specifieke (nauwe) taak. Voorbeelden hiervan zijn toepassingen voor het verbeteren van fraudedetectie bij schadeclaims, het voorspellen van klantvragen, of 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 beprijzing (premievaststelling) en acceptatie.

De 'voeding' voor AI-modellen, de inputdata, bestaat voornamelijk uit gegevens die intern bij verzekeraars beschikbaar zijn. Voor bepaalde processen wordt deze data ook meer en meer verrijkt met externe databronnen, zoals weerdatabases, data van Stichting CIS (Centraal Informatie Systeem) of data van de Kamer van Koophandel.

Aandacht voor technische aspecten van AI

De toepassing van AI-technieken in de verzekeringssector is nog volop in ontwikkeling. Van belang is dat verzekeraars, indien zij gebruik willen maken van AI, aandacht hebben voor de verschillende technische aspecten van AI-modellen, zeker waar deze nog niet volledig uitgekristalliseerd zijn.

Allereerst is het belangrijk dat verzekeraars vroegtijdig en systematisch nadenken over deze aspecten en vaststellen hoe zij AI wel en niet willen gebruiken. Een combinatie van breed in de organisatie gedragen kennis over AI en intern-ontwikkeld beleid over het gebruik ervan, verankerd in duidelijke governance, stelt verzekeraars in staat AI op verantwoorde wijze in te zetten en tevens kritische vragen bij die inzet te blijven stellen.

Meer specifiek verdienen drie technische onderdelen aandacht, namelijk de inputdata voor het model, het model/de techniek als zodanig en de modeluitkomsten en hoe hiermee om te gaan:

- Allereerst wordt de kwaliteit van een model, zeker bij AI-toepassingen, voor een groot deel bepaald door de inputdata waarmee het wordt gevoed. Voldoende aandacht voor de kwaliteit van inputdata en de mate waarin die data past bij de gewenste AI-toepassing is cruciaal. De kwaliteit van inputdata is temeer van belang wanneer – wat in toenemende mate gebeurt – verzekeraars voor het verrijken van datasets en het ontwikkelen van AI-toepassingen samenwerken met externe partijen. Hierbij verdient het zicht houden op de kwaliteit van de geïmporteerde data door de verzekeraar extra aandacht; de verzekeraar blijft immers verantwoordelijk voor gebruikte data en voor de AI-toepassing.
- Ten tweede moet voor iedere AI-toepassing een keuze worden gemaakt uit verschillende beschikbare AI-modellen en AI-technieken. In hoeverre berust die keuze op een bewuste afweging door de verzekeraar? Bij een bewuste afweging kan gedacht worden aan het tegen elkaar afzetten van voor- en nadelen van individuele technieken en modellen; en aan hoe die voor- en nadelen zich verhouden tot de beschikbare data en het proces waarin de AI-toepassing ingezet zal worden. Ook de vraag of AI-technieken en -modellen consistent

ingezet worden binnen de organisatie komt hier in beeld. Dezelfde afwegingen verdienen aandacht wanneer het gaat om extern-ontwikkelde en uitbestede AI-toepassingen.

- Tot slot is er de vraag hoe om te gaan met de resultaten van het model. Hier spelen vooral de vragen in hoeverre uitkomsten van een AI-toepassing herleid kunnen worden naar individuele inputparameters; en in hoeverre dit nodig is voor het proces waarin de AI-toepassing wordt ingezet? Daarnaast is het belangrijk dat verzekeraars nadenken over de wijze waarop en frequentie waarmee gevalideerd kan worden dat een AI-toepassing blijft doen waarvoor die ontwikkeld is.

Aandacht voor consumentengedrag en maatschappelijk acceptatie

Net als voldoende aandacht voor de technische aspecten van AI, is begrip van de maatschappelijke context waarbinnen AI wordt toegepast van groot belang. Zo is het onwenselijk wanneer AI-toepassingen klanten aanzetten in strijd te handelen met hun eigen (financieel) belang, bijvoorbeeld door in online beslisomgevingen op onredelijke wijze in te spelen op verwachte heuristieken en biases in consumentengedrag. Dergelijke inzichten zouden juist gebruikt moeten worden om een beslisomgeving te creëren die geschikte keuzes voor en door klanten stimuleert.

Daarnaast kunnen AI-toepassingen in de gebruikte data mogelijk nieuwe verbanden vinden die voorheen niet gelegd konden worden. Het is dan belangrijk dergelijke verbanden te toetsen op de mate waarin gebruik ervan maatschappelijk geaccepteerd is, of acceptabel zou zijn.

AI en solidariteit: kansen en risico's

De opkomst en ontwikkeling van AI in de verzekeringssector biedt mogelijkheden voor individuen die in het huidige systeem worden uitgesloten van verzekeringen, maar kan mogelijk ook druk leggen op de solidariteit tussen (groepen van) verzekerden. In deze afweging ligt er vooral een belangrijke rol voor de sector zelf, maar zowel de AFM als DNB levert graag een actieve bijdrage aan deze discussie, die overigens breder is dan alleen de mogelijke impact van gebruik van data en AI.

Aandachtspunten

Om bewustwording bij en dialoog met verzekeraars en andere stakeholders te stimuleren over het gebruik van AI in de verzekeringssector, hebben de AFM en DNB onderstaande punten geïdentificeerd als zijnde gebieden waarop verdere uitwisseling en verdieping van inzichten zinvol is. Deze aandachtspunten zijn gebaseerd op reeds gevoerde gesprekken met verzekeraars, stakeholders en experts en kunnen dienen als startpunt en structuur voor verdere dialoog:

Aandachtspunten

Inbedden van AI in de organisatie

1. Hoe kunnen verzekeraars beleid ontwikkelen omtrent het gebruik van AI?
2. Op welke wijze is dit beleid verankerd in de organisatie?

Technische aspecten van AI

3. Hoe kan de kwaliteit en volledigheid van gebruikte inputdata voor AI-toepassingen geborgd worden?
4. Hoe kunnen keuzes worden gemaakt omtrent het gebruik van specifieke AI-technieken en -modellen?
5. Welke mate van uitlegbaarheid van AI-toepassingen is gepast?
6. Hoe kan het gebruik van verbanden die leiden tot onwettig onderscheid (discriminatie) vermeden worden?
7. Welke governance en criteria zijn passend bij uitbesteding in het kader van AI-toepassingen?
8. Wat is een passende wijze van validatie voor AI modellen?

AI en de consument

9. Hoe kan worden voorkomen dat AI-toepassingen op onredelijke wijze inspelen op verwachte patronen of biases in consumentengedrag?
10. Hoe kunnen de uitkomsten van AI-toepassingen worden getoetst op maatschappelijke acceptatie?

1. Waarom deze verkenning?

De afgelopen jaren is de uitspraak "data is het nieuwe goud" meer en meer in zwang geraakt. Organisaties binnen en buiten de financiële sector hebben de beschikking over almaar meer data, die bovendien ook steeds gevarieerder wordt. De grotere beschikbaarheid van data, gecombineerd met een sterke toename in computerkracht, heeft ertoe geleid dat de ontwikkeling en toepassing van Artificiële Intelligentie (AI) de laatste jaren in een stroomversnelling zijn gekomen.

De verwachting is dat de inzet van AI ook de komende jaren zal blijven doorgroeien, zeker ook in de financiële sector.

De Nederlandse verzekeringssector vormt hierop geen uitzondering: ook hier is een toenemend gebruik van AI-toepassingen te zien én wordt verwacht dat de inzet van zulke toepassingen de komende jaren een verdere vlucht zal nemen.

Tegelijkertijd is privacy en de wijze waarop organisaties en bedrijven met data omgaan de laatste jaren hoger op de maatschappelijke en wetgevende agenda gekomen.

Ook leven discussies over de gewenste mate van solidariteit tussen individuen en groepen – en hoe solidariteit in verhouding staat tot toenemende data-analyse – breed in de samenleving.

De opkomst van AI-toepassingen raakt daarmee direct aan verschillende aspecten van het verzekeringsvak: het brengt als zodanig nieuwe kansen, maar ook risico's voor de verzekeringssector.

Definitie Artificiële Intelligentie in deze verkenning

Het begrip Artificiële Intelligentie (AI) wordt op verschillende wijzen gedefinieerd, met uiteenlopende betekenissen. Tegelijkertijd worden naast AI ook veel verschillende gerelateerde begrippen gebruikt wanneer verwezen wordt naar toepassingen die als doel hebben het sneller en efficiënter analyseren van meer en meer gevarieerde datapunten. Zo wordt in plaats van AI veel gesproken over machine learning, data science, data analytics, advanced analytics en Big Data & Artificial Intelligence (BDAI).

Deze verkenning gebruikt de termen AI, AI-modellen of AI-toepassingen en doelt hiermee op toepassingen die zijn gebaseerd op het analyseren van gevarieerde en grote hoeveelheden data met behulp van technieken als machine learning.

De technieken die worden gebruikt in AI-toepassingen (zoals machine learning) zijn 'intelligent' in de zin dat ze in staat zijn rationeel te optimaliseren: binnen een gegeven taak zijn ze in staat de beste actie te kiezen om een bepaald doel te bereiken, rekening houdend met gestelde criteria.

Naast het gebruik van intelligente technieken is ook data een voorwaarde voor en integraal onderdeel van AI-toepassingen. De ontwikkeling van AI wordt gevoed door de grotere volumes en diversiteit aan beschikbare data. In de verkenning zal het begrip 'big data' worden gebruikt om te verwijzen naar deze toename van inzetbare data.

Doel verkenning

In deze verkenning belichten de AFM en DNB de ontwikkelingen, kansen en risico's voor verzekeraars van het gebruik van AI. Dit gebeurt vanuit het perspectief van productontwikkeling, zorgplicht en integere en beheerste bedrijfsvoering. Doel van de verkenning is het starten van een dialoog met de sector en andere stakeholders. Deze dialoog kan worden gevoerd aan de hand van de 10 aandachtspunten die uit de verkenning naar voren zijn gekomen en in deze verkenning ook verder uiteen worden gezet. De aandachtspunten moeten worden gelezen als gebieden waarop verdere gedachte-uitwisseling en -verdieping nodig zijn om te komen tot een duidelijker en eenduidiger inzicht in hoe verzekeraars het best met AI kunnen omgaan.

Daar het gebruik van AI in de verzekeringssector nog volop in ontwikkeling is, is nu een goed moment om hierover een dialoog aan te gaan, mede om risico's in een vroegtijdig stadium te signaleren en te ondervangen. Deze verkenning dient er dan ook toe die dialoog verder aan te zwengelen en waar mogelijk te structureren.

Scope verkenning

Rondom AI doen veel verschillende termen de ronde en vaak hebben ze overlappende betekenissen. Deze verkenning zal ingaan op de inzet van AI in de verzekeringssector, waarbij AI wordt gedefinieerd als toepassingen en modellen voor het analyseren van gevarieerde en grote hoeveelheden data met behulp van technieken als machine learning.

De verkenning kijkt alleen naar AI-toepassingen bij processen die verzekerings-specifiek en verzekeringstechnisch van aard zijn, zoals het selecteren, inschatten en beprijzen van risico's, het behandelen van schades en het detecteren van mogelijk fraudeleuze claims. Er zal niet worden ingegaan op meer algemene manieren waarop AI ingezet kan worden, bijvoorbeeld in marketing- of in backoffice-processen die niet raken aan verzekeringstechnische processen als beprijzing of acceptatie. Dat zijn immers geen verzekeringsspecifieke toepassingen van AI. Ook zal niet worden stilgestaan bij de mogelijke impact van AI-toepassingen op de waardeketen of marktstructuur van de verzekeringssector.

Leeswijzer

De verkenning zal allereerst kijken naar de ontwikkelingen op het gebied van AI in de Nederlandse sector: hoe wordt AI momenteel gebruikt in verzekeringstechnische processen? De verkenning zal ingaan op de nieuwe soorten data die ingezet worden, maar ook op de toepassing van technieken als machine learning. Tevens zal worden uiteengezet hoe AI-toepassingen in de verzekeringssector zich naar verwachting in de komende jaren zullen ontwikkelen.

Vervolgens wordt ingezoomd op aandachtspunten en -gebieden voor het gebruik van AI. Hier kijkt de verkenning naar bijvoorbeeld de uitlegbaarheid en herleidbaarheid van uitkomsten van AI-toepassingen. Ook wordt belicht hoe de inzet van AI raakt aan governance, modelvalidatie, uitbesteding, het productontwikkelingsproces en de zorgplicht.

Tot slot wordt ingegaan op hoe toepassing van AI raakt aan de bredere discussie over solidariteit in de verzekeringssector.

2. Ontwikkelingen

Er wordt veel gesproken over AI-toepassingen en hoe deze de verzekeringssector gaan veranderen. Deze inzet van AI door verzekeraars is echter nog volop in ontwikkeling. Daarom kan allereerst de vraag worden gesteld welke toepassingen nu al worden gebruikt door Nederlandse verzekeraars en welke ontwikkelingen in de komende paar jaar verwacht kunnen worden.

In dit hoofdstuk zal allereerst worden ingegaan op de soorten AI-technieken en data die verzekeraars momenteel inzetten; en hoe zij verwachten dat die inzet in de komende jaren zal evolueren. Daarna wordt besproken in welke onderdelen van de waardeketen AI nu wordt toegepast en hoe die toepassing zich in de nabije toekomst zal ontwikkelen.

2.1 Gebruik van AI-technieken
(Welke AI-technieken worden gebruikt door Nederlandse verzekeraars?)

2.2 Gebruik van data
(Welke soorten data worden gebruikt door Nederlandse verzekeraars?)

2.3 Gebruik in de waardeketen
(In welke onderdelen van de waardeketen wordt AI toegepast?)

2.1 Gebruik van AI-technieken

Welke AI-technieken worden gebruikt door Nederlandse verzekeraars?

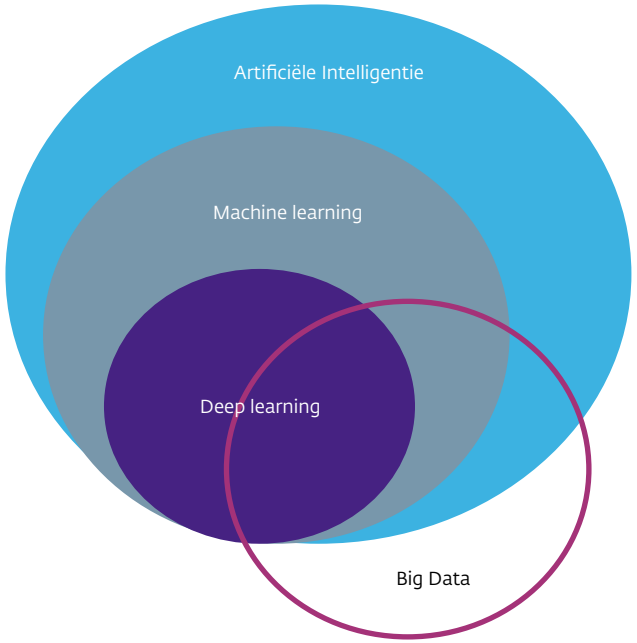
De laatste jaren komen meer en meer nieuwe technieken beschikbaar waarmee (grote hoeveelheden) gevarieerde data geanalyseerd kunnen worden. De belangrijkste van die technieken voor de verzekeringssector is machine learning, waarbij een systeem wordt getraind om een specifieke taak steeds optimaler uit te voeren. Deze training gebeurt niet aan de hand van expliciete instructies, maar aan de hand van optimalisatie-algoritmes en ingegeven data.

Tevens zien we de opkomst van AI-toepassingen die gebaseerd zijn op deep learning technieken: meer complexe machine learning technieken die gebruik maken van meerdere analyselagen (zie ook figuur 1).

Verschillende Nederlandse verzekeraars maken reeds gebruik van verschillende soorten machine learning toepassingen in verzekeringstechnische processen. Meest gebruikt hierbij zijn technieken als clustering, random forests, gradient boosting en deep neural networks. Inzet van machine learning toepassingen gebeurt vaak nog op een meer ad-hoc wijze en dergelijke toepassingen fungeren veelal als ondersteuning van, of als controle op, reguliere modellen.

Eén specifieke toepassing van machine learning betreft het gebruik van algoritmes om tekstdata beter te kunnen begrijpen en interpreteren; een toepassingsgebied dat bekend staat als Natural Language Processing (NLP). Verschillende Nederlandse verzekeraars passen NLP-technieken reeds toe, maar minder in verzekeringstechnische processen dan voor backoffice taken als het sorteren en alloceren van mails of post, of in het klantcontact door middel van virtuele assistenten.

Figuur 1 Verhouding AI-begrippen tot elkaar



In zijn algemeenheid kan daarmee worden gesteld dat voor verzekeringstechnische processen Nederlandse verzekeraars machine learning technieken nog niet op grote schaal of op structurele wijze inzetten. Veelal richt men zich voor dit soort taken en processen vooralsnog primair op het uitbreiden van bestaande statistische (regressie) modellen, bijvoorbeeld door nieuwe data en parameters aan deze modellen toe te voegen (zie ook 2.2 Gebruik van data).

2.1.1 Intern-gebouwde vs. extern-ingekochte modellen

Momenteel gebruiken Nederlandse verzekeraars voor hun AI-toepassingen (machine learning technieken) een mix van intern-ontwikkelde en extern-ingekochte ('off-the-shelf') modellen en platforms. Wel is er een beweging te zien naar het meer en meer intern bouwen van modellen, vooral naarmate de intern beschikbare expertise over machine learning groeit.

2.1.2 Manier van trainen van machine learning modellen

Machine learning modellen die worden toegepast zijn veelal getraind via *supervised* dan wel *unsupervised learning*. Er wordt weinig gebruik gemaakt van *reinforcement learning* – waarbij het model zichzelf aanpast op basis van een resultaat van een uitkomst – of van modellen die op continue basis worden geüpdatet aan de hand van nieuw beschikbare data. Verzekeraars (her)trainen hun modellen periodiek (wekelijks, maandelijks) met nieuw beschikbare data.

2.2 Gebruik van data

Welke soorten data worden gebruikt door Nederlandse verzekeraars?

In het kader van AI-toepassingen wordt vaak gesproken over de waarde en impact van 'big data'. Verzekeraars beschikken over grote volumes en diversiteit aan data, maar dit is niet vergelijkbaar met de 'big data' die processen bij Big Techs als Tencent, Google, Alibaba of Amazon genereren.

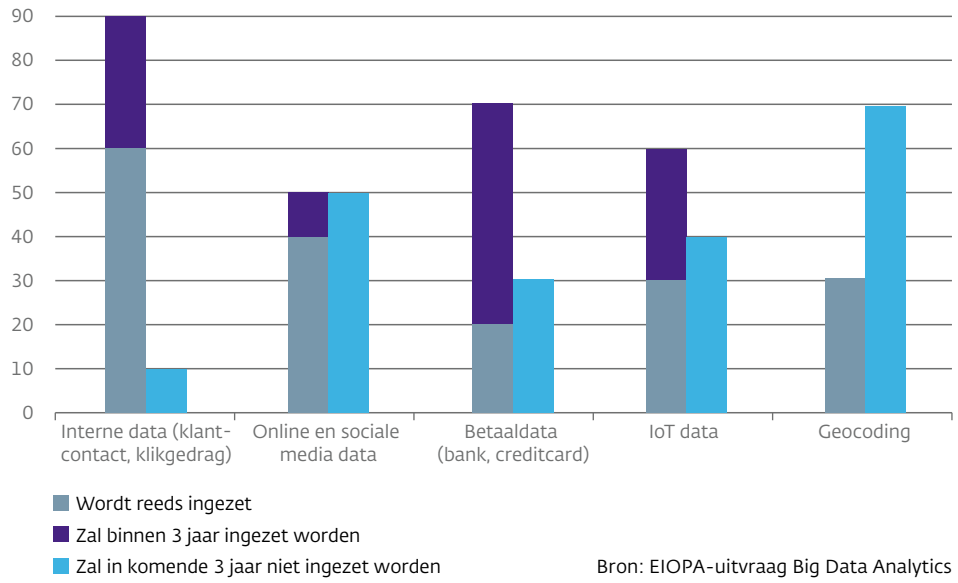
In de afgelopen jaren hebben verzekeraars zich dan ook vooral gericht op het beter ontsluiten van interne data (figuur 2), zowel gestructureerde (databases) als ongestructureerde data (tekst en scans). Dergelijke interne data wordt door verzekeraars gezien als het meest waardevol, vooral wanneer die data rechtstreeks input verstrekt over klanten of direct bruikbaar is in verzekeringstechnische processen. Voorbeeld hiervan is data over schades of over claimgedrag.

Naast interne gegevens wordt ook steeds meer gebruik gemaakt van externe databases; bijvoorbeeld data over betaalgedrag of kredietwaardigheid, maar ook data van de Kamer van Koophandel, Politie, UWV of CIS. Tevens wordt, vooral in traditionele schadesegmenten (auto, woon, landbouw), data uit weerdatabases of satellieten (geocoding, bodemdata) toegepast bij het beter inschatten of beheren van risico's. Zeker wanneer deze externe gegevens worden gecombineerd met interne data, worden ze gezien als potentieel zeer waardevol en verrijkend.

Op meer experimentele basis neemt de toepassing van Internet-of-Things (IoT) data toe. Het gaat hier vooral om auto-telematics, waarbij gegevens over rijgedrag en rijgewoontes (routes, tijden, etc.) geregistreerd worden. Grootschalige commerciële toepassing van IoT-data vindt in Nederland en andere landen (nog) niet plaats. Uitzondering hierop is het Verenigd Koninkrijk, waar IoT-data bijvoorbeeld wordt toegepast in autoverzekeringen voor hogere risicogroepen, zoals jongeren met weinig rijervaring. Als belangrijkste redenen voor de nog beperkte inzet van IoT-

Figuur 2 Gebruik nieuwe databronnen

% van deelnemende Nederlandse verzekeraars



Bron: EIOPA-uitvraag Big Data Analytics

data in de Nederlandse sector wordt gewezen naar de beperkte behoefte onder consumenten en zorgen over privacy, maar ook naar vragen over of via IoT-data momenteel voldoende context en inzicht verkregen kan worden in het te analyseren gedrag en verwante risico's. Ook bestaat in de sector de vraag of de data al in voldoende mate oorzakelijke verbanden weergeeft.

Tot slot is het gebruik van data van sociale mediaplatforms in verzekeringstechnische processen nog zeer beperkt. Enkele verzekeraars hebben testen gedaan, maar ervaringen zijn dat de meerwaarde van deze data vooralsnog klein is. Bovendien worden privacy-risico's van het gebruik van dit soort data als aanzienlijk beschouwd. Wel wordt door een aantal verzekeraars gekeken naar het gebruik van online-data, bijvoorbeeld data over het klikgedrag van consumenten op de website van de verzekeraar.

2.3 Gebruik in de waardeketen

In welke onderdelen van de waardeketen wordt AI toegepast?

AI-toepassingen kunnen ingezet worden in verschillende verzekeringstechnische processen, van productontwikkeling, risicoselectie, beprijzing en acceptatie, tot inschattingen van schadelast, claimsmanagement en fraudedetectie. De Europese verzekeringstoezichtautoriteit EIOPA heeft onlangs een uitgebreid onderzoek verricht, ook onder Nederlandse verzekeraars, naar de inzet van AI-toepassingen. Uit dit onderzoek – en uit gesprekken met verzekeraars – blijkt dat in het schade- en inkomenssegment Nederlandse verzekeraars de grootste mogelijkheden en impact van AI in de komende jaren verwachten op het gebied van beprijzing en acceptatie, als ook op het terrein van claimsmanagement en fraude (zie figuur 3). Voor zorgverzekeraars zal de impact van AI op het gebied van beprijzing en acceptatie waarschijnlijk een stuk kleiner zijn dan voor andere verzekeraars, omdat premiedifferentiatie en weigering van potentiële verzekerden, in ieder geval voor de basiszorgverzekering, niet is toegestaan.

2.3.1 Claimsmanagement & fraudebestrijding

Fraude

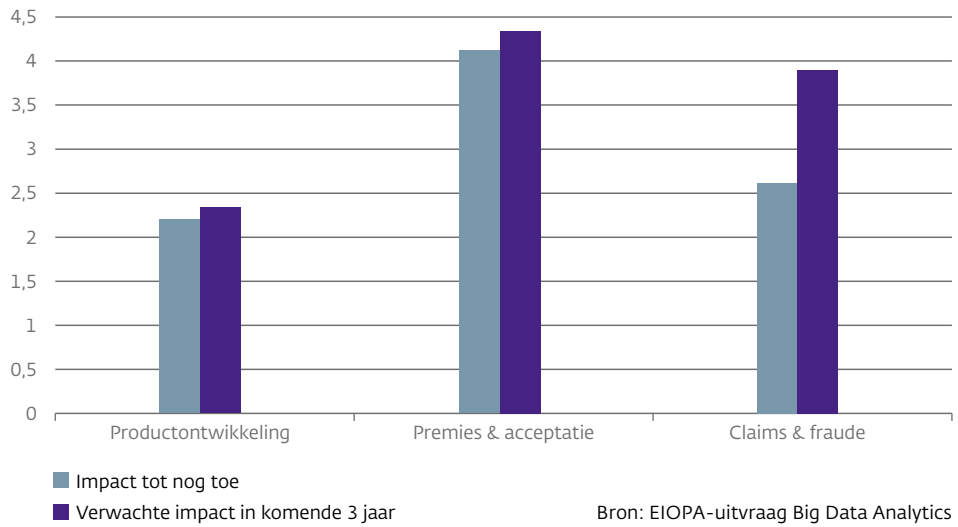
Voor fraudebestrijding wordt al op grotere schaal gebruik gemaakt van nieuwe databronnen – zoals foto's of prijslijsten van auto-onderdelen – en vaak ook van machine learning technieken, vooral anomaly detection (zie figuur 4).

Ook wordt claim scoring vaak gebruikt. Bij claim scoring krijgt iedere claim een score op basis van de karakteristieken van de claim. Komt de score boven een bepaalde drempelwaarde, dan wordt deze *geflagd* als mogelijk frauduleus. Claims die door het model als mogelijk frauduleus of afwijkend worden aangemerkt, worden vervolgens doorgaans door een schadebehandelaar beoordeeld. Volautomatische processen waarbij modellen beslissingen nemen over claims worden (nog) niet toegepast. Verzekeraars streven er wel naar om het percentage geautomatiseerde controle zo hoog mogelijk te krijgen. Hierbij worden verschillende aspecten meegewogen: de kosten van het mogelijk toestaan van *false negatives* (goedkeuring van claims die afgekeurd hadden moeten worden) als gevolg van geautomatiseerde afhandeling, de kosten van het handmatig controleren van mogelijke fraudeleuze claims om *false positives* (afkeuring van claims die eigenlijk goedgekeurd hadden moeten worden) te vermijden, als ook de reputatierisico's die aan *false positives* verbonden zijn.

Verzekeraars zien daarnaast sociaal-netwerkanalyse – waarbij het netwerk van bijvoorbeeld de persoon die de claim indient wordt geanalyseerd op aanwijzingen voor mogelijke fraude – als een mogelijk interessante AI-toepassing. Hetzelfde geldt

Figuur 3 Impact AI op activiteiten verzekeraars

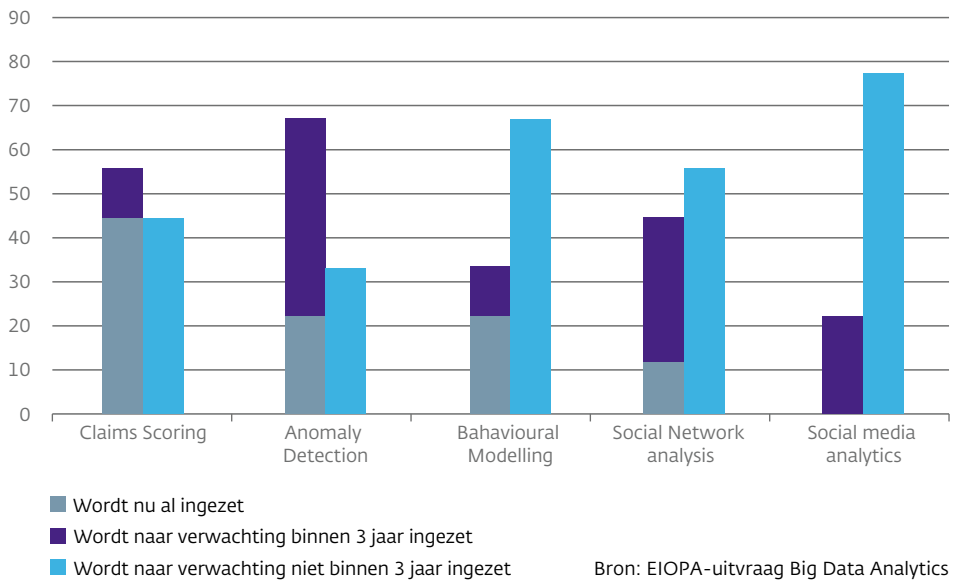
0 = lage impact, 5 = hoge impact



Bron: EIOPA-uitvraag Big Data Analytics

Figuur 4 Gebruik AI-toepassingen in fraudebestrijding

% van deelnemende Nederlandse verzekeraars



Bron: EIOPA-uitvraag Big Data Analytics

voor behavioural modelling, waarbij gedragingen van indieners van claims worden geanalyseerd. De verwachting is dat in de komende jaren dergelijke analyses meer en meer ingezet zullen gaan worden.

Het geautomatiseerd analyseren van sociale media, daarentegen, wordt vooralsnog niet gezien als een zeer bruikbaar of toepasbaar instrument voor fraudebestrijding, omdat verzekeraars de meerwaarde van sociale media data als relatief beperkt en de bijkomende privacyrisico's als relatief hoog inschatten.

Claimsmanagement

Los van fraudedetectie wordt, voor het afhandelen en managen van claims, hier en daar geprobeerd met behulp van AI claims nauwkeuriger te classificeren of karakteristieken van claims te voorspellen. Hiervoor wordt wijdverspreid gebruik gemaakt van externe databases met data over bijvoorbeeld de prijzen van reparaties van, of onderdelen voor, auto's. Een enkele verzekeraar zet ook in op het gebruik van machine learning technieken (bijvoorbeeld voor fotoherkenning of voor analyse van IoT-/sensordata) om de schadelast beter en sneller in te kunnen schatten. Van breed gebruik van AI-toepassingen op het terrein van claimsmanagement is echter nog geen sprake.

2.3.2 Beprijzing & acceptatie

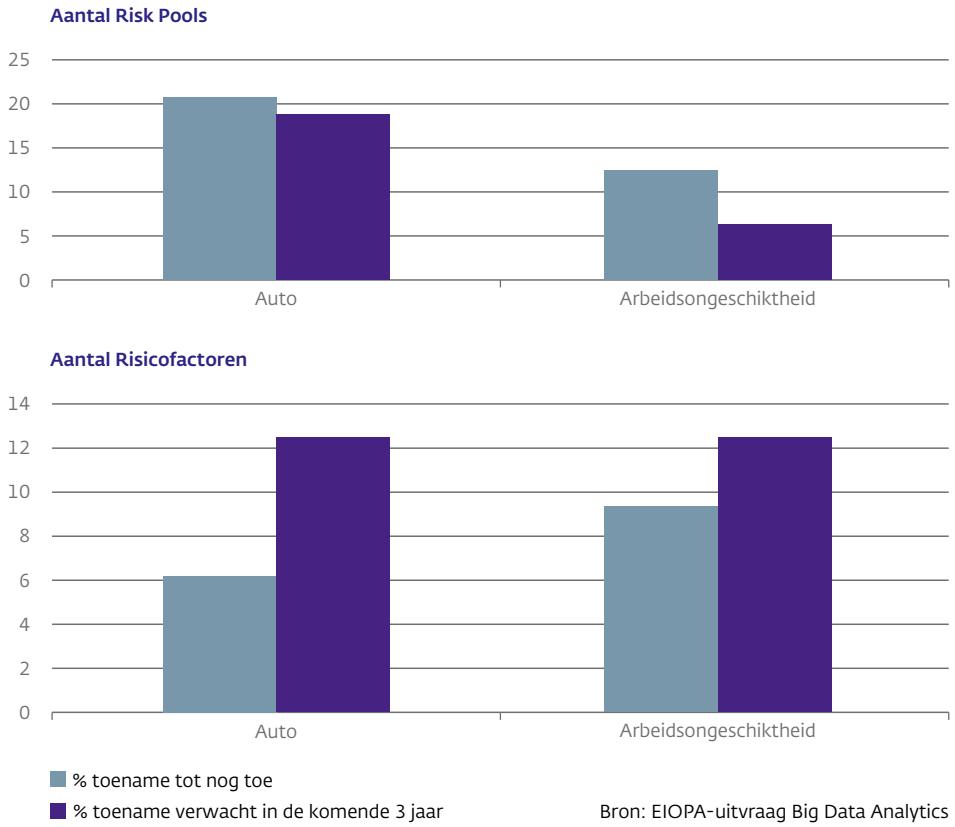
Risicobeprijzing

Op het gebied van het inschatten en beprijzen van risico's zijn de afgelopen jaren al – zij het beperkt – AI-toepassingen ingezet, bijvoorbeeld in het schadesegment (autoverzekeringen) maar ook, in mindere mate, in het inkomenssegment.

Zoals eerder aangegeven worden hier in de komende jaren verdere stappen verwacht. Zo zijn verschillende verzekeraars reeds bezig met het verfijnen van risico-inschattingen met behulp van aanvullende data. Deze aanvullende data bestaat uit intern-aanwezige data (vaak over bestaande klanten), openbaar-beschikbare data en extern-ingekochte data. Soms wordt de aanvullende data gebruikt ter versterking van bestaande risicofactoren, soms ook wordt data ingezet om nieuwe risicofactoren in modellen te introduceren of om verfijndere risicosegmenten te creëren (zie figuur 5). Deze nieuwe parameters zijn doorgaans van vrij traditionele aard: het gaat om parameters waarvoor intuïtief een oorzakelijk verband met schade bestaat.

Voor zowel autoverzekeringen als arbeidsongeschiktheidsverzekeringen is het aantal risicofactoren en risk pools dat Nederlandse verzekeraars toepassen bij het beprijzen toegenomen. De verwachting is dat deze trend ook de komende jaren zal voortzetten; verzekeraars verwachten vooral dat met het toenemend beschikbaar komen van data en data-analysetechnieken het aantal risicofactoren dat in modellen wordt gebruikt zal groeien.

Figuur 5 Aantal risicofactoren en riskpools neemt toe



Bron: EIOPA-uitvraag Big Data Analytics

Gedragsbeprijzing (dynamic pricing)

Bij gedragsbeprijzing – het premiedeel dat op basis van markomstandigheden en gedragsaspecten wordt vastgesteld; ook wel aangeduid als dynamic pricing of pricing optimisation – zijn meer AI-toepassingen zichtbaar. Gedragsbeprijzing wordt veelal gebaseerd op de loyaliteit (elasticiteit) van klanten, de lifetime value van klanten of de premies die worden gevraagd door concurrenten.

Dergelijke beprijzingsstrategieën zijn niet nieuw. Wel maken verzekeraars meer gebruik van nieuwe data – vaak een combinatie van 'nieuwe' interne data (bijvoorbeeld voor klantwaarde of loyaliteit, of door klik- en afsluitinformatie vanuit de website te analyseren) en externe databases – om gerichter aan dynamic pricing te kunnen doen. Pricing op basis van zeer specifieke en persoonlijke karakteristieken van (potentiële) klanten vindt (nog) niet plaats.

2.3.3 Technische voorzieningen

Er wordt niet tot weinig gebruik gemaakt van nieuwe technieken als machine learning voor de inschatting van de verwachte schadelast (technische voorzieningen); hiervoor worden veelal meer traditionele statistische modellen gebruikt. Enkele verzekeraars hebben aangegeven te kijken naar het toepassen van nieuwe technieken, vooral ter ondersteuning en toetsing van bestaande modellen.

2.3.4 Kapitaalmodellen

Momenteel zijn er in de Nederlandse sector geen toepassingen bekend van AI in modellen die de kapitaaleis of allocatie van kapitaal bepalen. De vraag is ook in hoeverre bijvoorbeeld machine learning technieken momenteel een grote meerwaarde kunnen hebben in dergelijke modellen. Van groot belang voor die modellen is om een 1:200 schokscenario te benaderen (de kans waarop de wettelijke kapitaaleis wordt bepaald); dit vraagt om het opstellen van een (kans)verdeling, terwijl machine learning technieken vooralsnog meer gericht zijn op het vinden van een beste schatting.

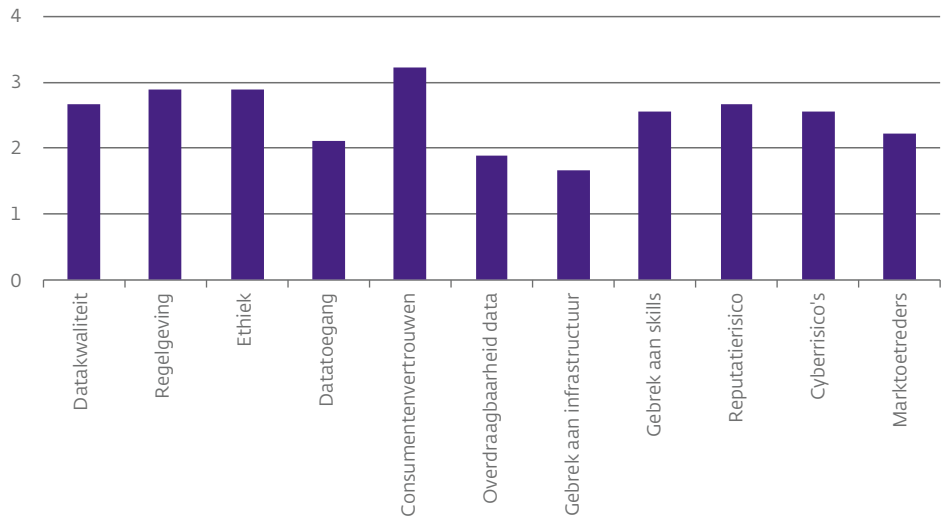
2.3.5 Customer experience & risicoselectie

Een belangrijke inspanning is in recente jaren geleverd bij het digitaliseren en verbeteren van de customer experience. Gedacht kan worden aan het inschatten van klantvragen door het analyseren van klikgedrag, text mining of NLP (mails, callcenters), verbeterde navigatie van websites, maar soms ook gepersonaliseerde website content. Deze toepassingen kunnen ook een rol spelen bij risicoselectie.

3. Aandachtspunten

Naast kansen brengt de inzet van AI ook op verschillende terreinen uitdagingen met zich mee. Verzekeraars zelf zien de risico's vooral op het terrein van de omgang met de consument: hoe behoudt men het vertrouwen van de consument, hoe kan AI op ethische wijze ingezet worden en wat zijn eventuele risico's voor de reputatie van een verzekeraar en de sector? Daarnaast zien verzekeraars ook regelgeving en de toezichthouder als een mogelijke bron van risico's: de vraag leeft hoe AI binnen de bestaande wet- en regelgeving kan worden toegepast (zie figuur 6).

Figuur 6 Uitdagingen voor het verantwoord implementeren van AI-toepassingen, volgens verzekeraars *(inschatting van de uitdaging door verzekeraars, 0 = geen uitdaging; 4 = zeer hoge uitdaging)*



Bron: EIOPA-uitvraag Big Data Analytics

Dit hoofdstuk zal ingaan op 10 punten omtrent de toepassing van AI door verzekeraars die verdere aandacht en dialoog verdienen. Deze aandachtspunten zijn onderverdeeld in 3 categorieën: inbedding van AI in de organisatie, technische aspecten van AI en AI en de consument.

3.1 Inbedden van AI in de organisatie
Hoe kan AI worden ingebed in de governance structuur en het beleid van verzekeraars?

3.2 Technische aspecten van AI
Welke technische aandachtspunten zijn er rond de ontwikkeling en toepassing van AI?

3.3 AI en de consument
Welke aandachtspunten liggen er rond de zorgplicht bij toepassing van AI?

3.1 Inbedden van AI in de organisatie

Hoe kan AI worden ingebed in de governance structuur en het beleid van verzekeraars?

Het risico bestaat dat verzekeraars onvoldoende systematisch nadenken over hoe ze AI willen toepassen. Er is momenteel veel aandacht voor AI. Dit brengt het risico met zich mee dat verzekeraars AI als een hype beschouwen en onvoldoende tijdig nadenken over of en hoe AI voor hen relevant is. Echter bestaat ook het omgekeerde risico, namelijk dat verzekeraars AI inzetten om 'ook mee te doen', zonder dat hieraan een heldere strategie ten grondslag ligt. Bij onvoldoende sturing richting data-analisten bestaat het risico dat op de werkvloer AI wordt toegepast op manieren die niet goed passen bij de *risk appetite* of strategie van de verzekeraar. Dit kan verzekeringstechnische risico's tot gevolg hebben, zoals het risico dat verkeerde keuzes worden gemaakt waar het gaat om acceptatie of premiestelling. Ook zijn er reputatierisico's voor de verzekeraar, bijvoorbeeld wanneer verzekerden worden onderworpen aan een gegevensanalyse die afwijkt van de verwachtingen die de verzekeraar hierover heeft gewekt.

Gegeven deze risico's is het van belang dat op bestuursniveau een helder beleid wordt vastgesteld over de inzet van AI-toepassingen (zie aandachtspunt 1). Onder de Solvency II-wetgeving [1] wordt van verzekeraars verwacht dat verantwoordelijkheden helder belegd zijn bij competente personen en dat (risico)strategie, doelen, risk appetite en risicolimieten, processen en rollen helder zijn vastgelegd. Risicobeleid over het gebruik van AI past hier ook in.

1 Artikelen 258 en 259 Solvency II Gedelegeerde Verordening

Aandachtspunt 1 – Bepalen van beleid voor AI-toepassingen

Het is van belang dat op bestuursniveau voldoende aandacht bestaat voor hoe de verzekeraar AI wel en niet wil toepassen. Het bestuur zou in die context bijvoorbeeld het antwoord willen weten op vragen als:

- Op basis van welke criteria besluit de verzekeraar over toepassing van AI-toepassingen (machine learning, big data)?
- Voor welke processen en onderdelen in de keten wil de verzekeraar wel – en voor welke niet – AI inzetten?
- Welke regels hanteert de verzekeraar voor het trainen en hertrainen van modellen? Met welke frequentie wil de verzekeraar haar modellen hertrainen? Hoe richt de verzekeraar de processen rondom (her)training in?
- Op het gebied van ethiek en maatschappelijke uitlegbaarheid:
  - Hoeveel differentiatie vindt de verzekeraar geoorloofd, zowel met betrekking tot risico-inschattingen als prijsoptimalisatie (dynamic pricing, gedragspricing)?
  - Op basis van welk soort inputdata wil de verzekeraar differentiatie toepassen?
  - In hoeverre wil de verzekeraar AI inzetten op een wijze die verzekerden de kans biedt hun risicoprofiel te verbeteren (preventie in zorg of in de woning, rijgedrag)?
  - In hoeverre wenst de verzekeraar premiekorting te geven in ruil voor het afgeven van data?
- Hoe belegt de verzekeraar de verantwoordelijkheden, processen en rollen bij het maken van beslissingen binnen de gestelde beleidskaders?
- Hoe waarborgt de verzekeraar dat de verantwoordelijkheid voor AI-toepassingen duidelijk is belegd binnen het bestuur en dat het verantwoordelijke bestuurslid voldoende technische kennis en afdoende ervaring heeft om de risico's van AI-toepassingen in te schatten, te toetsen en sturing te geven richting de organisatie?

Onvoldoende kennisdeling en toetsing leidt binnen de organisatie tot risico's voor een goede implementatie van AI. Het opstellen van beleid over hoe en op welke manier de verzekeraar AI in wil zetten is noodzakelijk maar niet voldoende: wanneer de organisatie onvoldoende op de hoogte is van wat het beleid is en wanneer de relevante onderdelen van de verzekeraar – zoals de data science afdeling(en), actuariële functie, risk management, IT – onvoldoende met elkaar communiceren en elkaar uitdagen, blijft het risico bestaan dat de manier waarop AI wordt gebruikt niet in lijn is met de *risk appetite* van de verzekeraar. Dit kan er alsnog toe leiden dat klanten en de verzekeraar zelf zijn blootgesteld aan onvoorziene verzekeringstechnische en operationele risico's. Dergelijke risico's worden verder vergroot wanneer kennis of inzet van AI teveel versnipperd zijn binnen de organisatie. Dit kan gebeuren wanneer verschillende onderdelen van de verzekeraar of de productieketen ieder op eigen wijze AI toepassen zonder een overkoepelend intern toezicht, bijvoorbeeld via een intern overkoepelend data science team. Het is van belang dat kennis van AI in de organisatie goed wordt geborgd door middel van een duidelijke governance. Zo kan ervoor worden gezorgd dat (uitkomsten van) modellen breed gevalideerd kunnen worden en dat continuïteit gewaarborgd is. Dit kan worden bewerkstelligd op verschillende manieren, bijvoorbeeld door centralisatie van data science expertise binnen een specifiek team, of door middel van structurele en gestructureerde deling van dergelijke kennis binnen de verzekeraar.

Binnen Solvency II is het concept van elkaar uitdagen (*challenge*) vastgelegd in de rollen van sleutelfuncties, waarbij vooral de Risicomanagementfunctie (model-validatie) en de Actuariële functie (gebruikte methodes, datakwaliteit) directe challenge bieden. De interne audit functie heeft op haar beurt de taak de andere sleutelfuncties onafhankelijk te controleren. [2] Deze functies kunnen hun respectievelijke rollen ook oppakken waar het gaat om AI-toepassingen. Gegeven de technische complexiteit, de relatief hogere frequentie waarmee AI-modellen worden aangepast en het feit dat het gebruik van AI door verzekeraars nog relatief nieuw is, is het aan te bevelen om de challenge zoveel mogelijk plaats te laten vinden op het niveau waar AI-toepassingen worden ontworpen en ingezet.

Aandachtspunt 2 – Communicatie over en verankering van AI-toepassingen

Een heldere interne communicatie over wat het bestuur wel en niet wil met betrekking tot AI-toepassingen geeft sturing aan de inzet van AI. Daarnaast is gestructureerde deling van kennis over en ervaring met AI van belang voor het opbouwen van institutionele kennis over AI. Doorlopende challenge op AI-toepassingen met en tussen verantwoordelijke bestuursleden en relevante sleutelfuncties vergroot verder de aandacht voor het mitigeren van risico's van AI-toepassingen.

2 Artikelen 269, 271, 272 Solvency II Gedelegeerde Verordening

3.2 Technische aspecten van AI

Welke technische aandachtspunten zijn er rond de ontwikkeling en toepassing van AI?

3.2.1 Voorafgaand aan training: datakwaliteit en modelkeuze

Het beschikken over kwalitatief goede en gevarieerde inputdata is een voorwaarde voor het toepassen van AI. De kalibratie van AI-toepassingen, vooral machine learning modellen, is volledig gebaseerd op het optimaal analyseren van verbanden zoals die in de data bestaan. Feitelijk correcte en representatieve data is daarom cruciaal voor het kalibreren van machine learning toepassingen. Bovendien geldt dat hoe gecompliceerder en minder doorzichtig modellen worden, hoe lastiger het wordt om met behulp van sanity checks fouten of biases in de data te traceren en verhelpen.

Ervaringen uit de praktijk wijzen erop dat, ondanks het werk dat verzekeraars in recente jaren hebben gedaan om de kwaliteit en ontsluiting van data te verbeteren, het volume, de volledigheid en de kwaliteit van data soms nog (aanzienlijk) tekortschiet om machine learning op brede schaal effectief in te zetten. Bijzondere aandacht moet hierbij uitgaan naar data over individuele klanten, o.a. door het gebruik van volmachten of tussenpersonen is zulke informatie nu vaak nog niet beschikbaar of incorrect.

Gebrek aan voldoende, correcte, volledige of gevarieerde data kan aanleiding geven tot verzekeringstechnische risico's. Dergelijke 'vuile' data kan leiden tot opname van foutieve verbanden in het model. Daarnaast geldt zeker voor machine learning modellen –waar de kalibratie en vorm van het model wordt geregisseerd door de inputdata –dat verbanden (risico's) die niet in de data zitten ook niet door het model herkend zullen worden. Dit risico wordt groter naar mate de data minder divers is. Foutieve data kan daarnaast ook risico's opleveren voor verzekerden, wanneer zulke data leidt tot verkeerde beslissingen, bijvoorbeeld waar het gaat om premievaststelling, acceptatie of keuze voor een product dat in werkelijkheid niet passend is.

Aandachtspunt 3 – In-control zijn over inputdata

Alvorens AI toe te passen, is het van belang dat de verzekeraar aan kan tonen dat zij in control is over de relevante inputdata. De volgende vragen verdienen daarbij aandacht:

- Heeft een verzekeraar een up-to-date overzicht van de te gebruiken data-elementen?
- Zijn datakwaliteitsnormen opgesteld voor de te gebruiken inputdata?
- Wordt de datakwaliteit van inputdata op continue wijze gemonitord aan de hand van controls?
- Is een risicobeoordeling gemaakt van de kwaliteit en volledigheid van de inputdata?
- Wordt adequaat ingespeeld op eventuele tekortkomingen in de data ten opzichte van gestelde normen?
- Voldoet de data die de verzekeraar als input gebruikt aan de door de verzekeraar gestelde datakwaliteitsnormen?

Naast het ontbreken van materiële fouten in de te gebruiken inputdata, is het van belang dat de gebruikte datasets zover mogelijk ontdaan zijn van ongewenste biases en aannames.

Om deze risico's effectief te beheersen is het belangrijk om in-control te zijn over de 'sanitaire' kwaliteit van de data: de te gebruiken datasets moeten met andere woorden vrij zijn van materiële fouten of missende datapunten. Om dit te waarborgen zouden verzekeraars de processen die zijn omschreven in de DNB Guidance over datakwaliteit voor Solvency II rapportageprocessen[3] ook kunnen inzetten op hun AI-toepassingen. Bij de challenge op de kwaliteit van data gebruikt in AI-toepassingen ligt tevens een belangrijke rol voor sleutelfuncties, vooral de Actuariële functie.

Bij het beoordelen van de geschiktheid van inputdata voor het gebruik in machine learning modellen moet echter verder worden gekeken dan of de data feitelijk correct en compleet is. Ook is van belang dat de data afdoende divers en representatief is voor de toepassing waarvoor de data wordt ingezet. Hierbij komen vragen aan orde als: zijn er afdoende ongebruikelijke waarden (outliers) opgenomen in de dataset? Welke inputparameters zijn wel en niet opgenomen in de dataset en om welke reden? Om representativiteit te verhogen kan ook gebruik worden gemaakt van bijvoorbeeld scenario-analyse en simulatie van data. Ook verdient het bestaan en waar mogelijk tegengaan van eventuele biases en aannames in de data an sich aandacht (zie ook aandachtspunt 6).

Bij de keuze voor het gebruik van bepaalde AI toepassingen is het van belang dat verzekeraars zowel rekening houden met statistische risico's als de mogelijke complexiteit van de toepassingen. Dit geldt vooral voor machine learning modellen. Statistische risico's manifesteren zich bijvoorbeeld wanneer een verzekeraar ervoor kiest een machine learning model toe te passen waarvoor de verzekeraar onvoldoende (diverse) data heeft. Ook wanneer de data correct en volledig is, kan dan bijvoorbeeld het risico op een overfit toenemen, waardoor het model foutieve uitkomsten geeft bij verzekerden met kenmerken die afwijken van die in de gebruikte training set. Een ander risico bij keuzes van modellen betreft de complexiteit van het model: wanneer de verzekeraar gebruik maakt van modellen waar het moeilijker is de verbanden in het model (op intuïtieve wijze) te analyseren, loopt de verzekeraar het risico dat het de uitkomsten en verbanden niet goed kan overzien en daarmee mogelijk niet langer voldoende in-control is over de modellen. Dit kan in het ergste geval leiden tot verkeerde beslissingen, maar ook tot discriminatie. Zeker bij processen die een directe impact hebben op verzekerden, is het van belang om bij de modelkeuzes rekening te houden met de complexiteit. Waar van toepassing kan het zelfs noodzakelijk zijn om een expliciete afweging te maken tussen de effectiviteit en de complexiteit van modellen.

3 Zie DNB (2017) 'Guidance beheersing Solvency II datakwaliteit voor verzekeraars', www.toezicht.dnb.nl/binaries/50-236703.pdf

Aandachtspunt 4 – Modelkeuze

Bij de keuze van een AI-/machine learning model verdienen de volgende vragen aandacht:

- Kan de verzekeraar op systematische wijze onderbouwen waarom voor een bepaald model en een bepaalde techniek is gekozen?
- Is de modelkeuze gebaseerd op de kwantiteit, kwaliteit en diversiteit van de beschikbare inputdata?
- Is bij de keuze voor een model of techniek ook gekeken naar factoren als uitlegbaarheid, complexiteit, betrouwbaarheid, en niet alleen naar best fit?
- Bestaat een zekere mate van consistentie tussen de modellen en technieken die worden gebruikt voor het vaststellen van premies en de modellen en technieken die worden gebruikt voor het vaststellen van technische voorzieningen?
- Zijn bij de keuze voor een bepaald model experts uit de relevante (business) domeinen betrokken, inclusief domeinen als IT, Actuariële functie en Risicomanagementfunctie (modelvalidatie)?
- Kan de verzekeraar inzicht geven in hoe de gekozen techniek in meer algemene zin werkt, voor welke typen processen of typen datasets een bepaalde techniek meer geschikt is en voor welke minder?
- Kan de verzekeraar inzicht geven in de omstandigheden waaronder gebruik van de gekozen techniek niet (langer) passend zou zijn? Hoe wordt hier periodiek op getoetst?

3.2.2 Omgang met uitkomsten

Zeker voor het gebruik in gevoelige processen is uitlegbaarheid van uitkomsten van AI-toepassingen van groot belang. Uitlegbaarheid betekent allereerst dat een verzekeraar in staat is aan te geven hoe bepaalde inputdata leidt tot een bepaalde uitkomst. Dit komt erop neer dat kan worden aangegeven welke parameters in belangrijke mate hebben bijgedragen aan de uitkomst, als ook een kwantificering van die bijdrage. Daarnaast bedoelen we met uitlegbaarheid dat kan worden aangegeven welke veranderingen in individuele inputwaarden nodig zijn om een bepaalde verandering in de uitkomst van het model tot stand te brengen.

Voor sommige machine learning technieken is een dergelijke mate van uitlegbaarheid lastig of zelfs onmogelijk. Modellen die met dergelijke technieken worden opgebouwd, worden doorgaans black box-modellen genoemd. Enkele voorbeelden van black box technieken zijn deep neural networks en, in mindere mate en afhankelijk van hoe het model is opgebouwd, random forests. Hoewel er verschillende mogelijkheden zijn om verbanden binnen black box-modellen enigszins bloot te leggen (bijvoorbeeld via explainer modellen of partial dependence analyse), gelden dergelijke observaties vaak alleen voor één bepaalde combinatie van input en uitkomst. Dit is een belangrijk verschil met white box-modellen, waar in principe voor alle inputverbanden het verband tussen inputwaarden en uitkomsten – als ook de statistische betrouwbaarheid van dat verband – kan worden afgelezen.

Het dilemma van uitlegbaarheid en black box-modellen komt nadrukkelijker naar voren met de invoering van nieuwe Europese Verordening ten aanzien van gegevensbescherming (AVG), die rechten voor consumenten versterkt wanneer zij onderworpen worden aan geautomatiseerde besluitvorming. Zo moeten in dit geval verzekeraars inzicht kunnen geven aan verzekerden in de onderliggende logica van de modellen die worden gebruikt in de besluitvorming en in de gevolgen die

geautomatiseerde besluitvorming heeft voor de verzekerde. Ook kan de consument bezwaar maken tegen het genomen besluit.[4]

Aandachtspunt 5 – Uitlegbaarheid uitkomsten

Bij het bepalen in hoeverre een model of machine learning techniek passend is voor een proces met geautomatiseerde besluitvorming, verdienen de volgende vragen aandacht:

- In welke mate worden de modellen ingezet in processen die een directe en grote impact hebben op klanten – en daarmee eventueel risico's voor de verzekeraar meebrengen met betrekking tot productontwikkeling, zorgplicht, juridische en reputatierisico's – of in processen die een directe en grote impact hebben op de stabiliteit van de verzekeraar?
- In welke mate is het model of die machine learning techniek uitlegbaar? Met andere woorden: in welke mate is het mogelijk verbanden tussen inputparameters en modelresultaten te achterhalen?
- Welke mate van uitlegbaarheid is passend voor dat specifieke proces?

Uitlegbaarheid kan het best op een meer case-by-case basis worden bekeken, omdat niet voor elke machine learning techniek de mate van uitlegbaarheid eenzelfde is en tevens niet ieder proces vraagt om eenzelfde mate van uitlegbaarheid. Hierbij hoeven black box modellen dus niet per se in de ban te worden gedaan, maar moet wel goed worden gekeken naar de mate van uitlegbaarheid die gewenst is voor een proces; en hoe die mate van uitlegbaarheid kan worden bewerkstelligd. Hieronder worden enkele voorbeelden behandeld:

Back-office processen: bij dergelijke processen kan worden gedacht aan het sorteren van post of van mail, of optimaliseren van de inzet van callcenters. Hierbij is de impact op de klant – ook waar het gaat om discriminatoire verbanden – vaak beperkt. Inzet van minder uitlegbare technieken en modellen heeft hier een kleine impact op de klant en op de stabiliteit van de verzekeraar. Eventuele gebreken in 'fit-for-purpose' van modellen kunnen aan de hand van outputchecks worden opgespoord en geadresseerd. Daarom is in dergelijke processen het gebruik van black box-modellen acceptabeler dan voor andere processen.

Modellen voor beprijzing, acceptatie en fraudedetectie: dit soort modellen heeft een directe impact op de klant, temeer wanneer deze modellen gebruikt worden in geautomatiseerde besluitvorming. Zorgen omtrent productontwikkeling, zorgplicht, uitlegbaarheid en discriminatie zijn in beginsel groot, als ook de hieraan gekoppelde reputatie- en juridische risico's voor de verzekeraar. Daarmee is het van belang dat voor individuele inputparameters aangegeven kan worden of en in welke mate deze bij hebben gedragen aan de uitkomst van het model en welke veranderingen in inputparameters nodig zijn om een verandering in de modeluitkomst te bewerkstelligen. Een black box-model, waarbij zulke uitleg niet of moeilijk gegeven kan worden, zal naar verwachting dan ook niet goed passen in een proces waarin geautomatiseerde besluitvorming plaatsvindt die de klant rechtstreeks raakt.

Wel kunnen black box-technieken worden gebruikt ter ondersteuning van besluitvorming door andere modellen of bijvoorbeeld door frauderechercheurs of 1e-lijns actuarissen. Hierbij moet een afweging worden gemaakt omtrent de mate waarin de verzekeraar *false positives* en *false negatives* acceptabel vindt. Ook is het van belang dat de 'menselijke tussenkomst' geen formaliteit wordt.

4 Artikelen 15(1)(h) en 21 AVG.

Modellen ten behoeve van berekening technische voorzieningen of kapitaaleisen: voor dergelijke modellen zijn zorgen omtrent discriminatie of het direct en onterecht benadelen van consumenten van kleinere omvang dan bij premievaststelling of acceptatie. Ook zou in principe een kleinere mate van uitlegbaarheid van het model mogelijk zijn dan bijvoorbeeld in processen als beprijzing of acceptatie. Wel is het vanuit prudentieel oogpunt van belang dat de verzekeraar een goed inzicht heeft in de statistische betrouwbaarheid en de onzekerheidsmarge van uitkomsten. Inzet van een volledige 'black box' voor automatische besluitvorming in zulke processen zou dan ook veelal niet wenselijk zijn. Het is van het grootste belang dat kan worden aangetoond dat de modeluitkomsten correcte inschattingen geven van de benodigde voorzieningen. Dit kan deels worden bereikt in het validatieproces, bijvoorbeeld door back-testing en het testen van voorspellingen van het model versus gerealiseerde schade. Echter dient een zekere mate van uitlegbaarheid behouden te blijven, ook om effectieve challenge door de Actuariële Functie mogelijk te maken. Dit is voor segmenten met langere looptijden, bijvoorbeeld levensproducten, van groter belang, omdat het hier vaak langer duurt voordat gerealiseerde schadelast bekend wordt waarmee modeluitkomsten vergeleken kunnen worden.

Wanneer discriminatoire biases in AI-toepassingen niet goed vermeden kunnen worden, zou moeten worden overwogen om de toepassing niet in te zetten. Discriminatie is het op onwettige wijze anders behandelen, achterstellen of uitsluiten van mensen op basis van (persoonlijke) kenmerken; zogenaamde verboden discriminatiegronden. Wanneer verzekeraars verzekerden analyseren (bijvoorbeeld in het acceptatieproces, het vaststellen van premies of eventueel als onderdeel van fraudedetectie) bestaat een risico dat als onderdeel van AI-toepassingen verbanden en parameters worden toegepast die discriminatoir zijn.

Aandachtspunt 6 – Vermijden gebruik verbanden die leiden tot onwettig onderscheid (discriminatie)

De verzekeraar dient systemen en processen ingericht te hebben die voorkomen dat AI-toepassingen leiden tot discriminatoire uitkomsten.

De volgende vragen verdienen aandacht bij het opstellen van dergelijke systemen en processen waar het gaat om AI-toepassingen:

- Hoe kan challenge op mogelijke discriminatoire bias in de inputvariabelen worden ingericht?
- Hoe kunnen uitkomsten worden getest op discriminatoire bias? Hierbij kan worden gedacht aan het gebruik van adversarial modelling, als ook aan het uitvoeren van steekproeven met identieke testgroepen, waarbij alleen een discriminerende (proxy)variabele verschilt tussen groepen. Indien significante verschillen uit het model komen, kan dit duiden op mogelijke discriminatoire bias in het model. Hierbij kan ook gebruik gemaakt worden van gefabriceerde casussen.
- Hoe kunnen tests op discriminatoire bias verfijnder en robuuster worden gemaakt? Hierbij kan worden gedacht aan het specifiek testen van biases in de false positive uitkomsten (in plaats van het alleen testen van de overall modeluitkomsten).

Indien voor een AI-model niet voldoende kan worden vastgesteld dat het geen ongeoorloofde discriminatoire biases bevat, moet worden overwogen of het prudent is het model toe te passen in processen die direct raken aan klanten (bijvoorbeeld pricing en acceptatie, fraudedetectie).

De wet bepaalt de gronden waarop onder meer verzekeraars géén onderscheid mogen maken. Voorbeelden hiervan zijn geslacht, etniciteit, geloof, geaardheid of het hebben van een handicap.[5] Directe toepassing van zulke parameters is relatief gemakkelijk te vermijden, maar van belang is ook om indirecte toepassing van – proxies voor – discriminatoire variabelen te voorkomen.

Dit is geen nieuw risico; immers kunnen ook minder uitgebreide of gecompliceerde modellen discriminatoire proxies bevatten. Wel kunnen binnen AI-toepassingen – en vooral binnen minder uitlegbare machine learning technieken – meer en minder zichtbare verbanden worden gelegd die mogelijk discriminatoir zijn en die moeilijker te traceren zijn.

Het is belangrijk dat verzekeraars een proces opstellen waardoor zij op continue wijze de inputparameters en mogelijke discriminatoire verbanden in modellen kunnen challengen. Zeker gegeven de mogelijke gevolgen die discriminatoire biases kunnen hebben voor klanten – en de gerelateerde juridische en reputatierisico's voor een verzekeraar – dient een verzekeraar een model alleen voor processen in te zetten indien met voldoende zekerheid kan worden vastgesteld dat het niet leidt tot verboden discriminatoire uitkomsten.

Indien een verzekeraar toch een proxy voor een verdacht of onwettelijk onderscheid wil gebruiken, dient hiervoor gekeken te worden naar een onderbouwing volgens de gangbare juridische testen: zo moet een objectieve doelstelling kunnen worden gegeven voor het onderscheid en moet de verzekeraar kunnen aantonen dat het gebruik van het onderscheid proportioneel is om de doelstelling te bereiken. Ook zou de variabele/proxy voldoende afgebakend moeten kunnen worden. Zeker bij minder uitlegbare (machine learning) modellen kan het lastig zijn de precieze verbanden die in het model gebruikt worden te achterhalen, waardoor het ook moeilijker wordt de juridische hordes te nemen.

Tot slot is het van belang dat verzekeraars hun bias-analyse toespitsen op het doel waarvoor het model wordt ingezet. Zo is het bij het vaststellen van technische voorzieningenprocessen minder problematisch of juist wenselijk om een bepaald onderscheid (bijvoorbeeld man/vrouw) te maken, terwijl dit bij het vaststellen van premies mogelijk zeer onwenselijk en mogelijk onwettig is.

3.2.3 Gebruik van externe data en modellen

Bij uitbesteding van AI-toepassingen is het goed monitoren, toetsen en challengen van uitbestede toepassingen cruciaal. Verzekeraars maken veel gebruik van externe expertise in de vorm van data of algoritmes/modellen. Het betreft veelal het inkopen van externe databases, het samenwerken met externe consultants voor het trainen van machine learning modellen, of het extern inkopen van vooraf getrainde modellen.

Zonder goede uitbestedingsprocessen kan dit leiden tot het risico dat verzekeraars minder inzicht hebben in de kwaliteit van de datasets die wordt geleverd, of in de wijze waarop externe modellen zijn getraind en hoe deze functioneren. Zeker wanneer kritische of belangrijke verrichtingen worden uitbesteed is er daarmee a priori een risico dat uitbesteding leidt tot vergroting van verzekeringstechnische en operationele risico's

Met het toenemend gebruik van externe data en modellen is het van steeds groter belang dat verzekeraars scherp kijken naar hoe zij met externe partijen samenwerken op het gebied van data en data-analyse. Dit verdient aandacht in

5 Artikel 5 AWGB en Artikel 9(1) AVG

Aandachtspunt 7 – Uitbesteding van AI-toepassingen

Het is van belang dat verzekeraars zicht hebben op AI-toepassingen, ongeacht of deze intern zijn ontwikkeld of zijn uitbesteed aan een externe partij.

Verzekeraars moeten hun uitbestedingen en samenwerkingen met externe partijen (bijvoorbeeld met betrekking tot data en modellen) vergelijken met de door DNB opgestelde Good Practice op het gebied van uitbesteding (2018). Op relevante processen kan dan de Guidance beheersing Solvency II datakwaliteit door verzekeraars (2017) toegepast worden.

Verzekeraars dienen een uitbestedingsbeleid te hebben opgesteld en te bepalen – aan de hand van de criteria besproken in de DNB Good Practice – welke uit te besteden AI-toepassingen kritiek zijn. Ook dienen verzekeraars een proces te hebben voor het monitoren van uitbestede processen. Hierbij verdienen de volgende aandacht:

- Heeft de verzekeraar voldoende kennis om inzicht te hebben in de werking van de externe toepassing?
- Zijn er afspraken gemaakt met externe partijen met betrekking tot de kwaliteit en herkomst van de data die de externe partij levert en de wijze waarop externe modellen zijn getraind/gekalibreerd?

Wanneer bovenstaande vragen niet afdoende zijn ondervangen met betrekking tot een bepaalde externe partij, moet worden overwogen of samenwerking met die externe partij wenselijk is.

Verzekeraars zouden daarnaast ook individueel dan wel waar mogelijk en gepast collectief, stappen kunnen nemen om te voorkomen dat uitbesteding van AI-toepassingen continuïteit van de onderneming ondermijnt. Onderdeel hiervan is het maken van afspraken met externe partijen over beschikbaarheid van toepassingen en data, als ook het voorkomen van een te grote afhankelijkheid van een of meerdere externe partijen.

het uitbestedingsbeleid, ook specifiek gericht op de scope van uitbesteding: welke activiteiten wil de verzekeraar uitbesteden en welke niet? Van welke externe databronnen en toepassingen willen verzekeraars gebruik maken en met welk doel?[6]

Volgens Solvency II gelden extra criteria voor kritieke uitbestedingsprocessen. Om vast te stellen welke uit te besteden AI-activiteiten kritiek zijn, zou de verzekeraar zich de vragen moeten stellen die uiteengezet worden in DNB *Good Practice Uitbesteding Verzekeraars:*

- Is de activiteit inherent kritiek voor de verzekeraars, bijvoorbeeld is de activiteit kritiek in het kunnen nakomen van verplichtingen aan polishouders?
- Wat zijn de operationele effecten (reputatie, juridisch) bij onderbrekingen?
- Welke impact kunnen verstoringen in de uitbesteding hebben voor inkomsten van de verzekeraars?
- Welke impact heeft een schending van vertrouwelijkheid bij een uitbesteding op verzekeraars en hun polishouders?

Volgens Solvency II moeten verzekeraars uitbestedingsovereenkomsten opstellen met externe partijen[7]. Dit geldt ook bij uitbesteding van AI-toepassingen. In deze overeenkomst dienen naast de wettelijke verplichtingen ook de verwachtingen wat betreft de te leveren kwaliteit van AI-toepassingen vastgelegd te worden. Bij uitbesteding van dataverzameling moet bijvoorbeeld via een dataleveringscontract afspraken worden gemaakt over de te leveren datakwaliteit.[8]

6 Zie ook DNB (2018), 'Good Practice Uitbesteding Verzekeraars', https://www.toezicht.dnb.nl/2/50-237170.jsp
7 Artikel 274 Solvency II Gedelegeerde Verordening
8 Zie ook DNB (2017) 'Guidance beheersing Solvency II datakwaliteit voor verzekeraars' www.toezicht.dnb.nl/binaries/50-236703.pdf

3.2.4 Validatie

Het is van belang dat validatieprocessen zo ingericht zijn dat kan worden vastgesteld of AI-toepassingen *fit for purpose* zijn, ook wanneer die frequent of continu geüpdatet of hertraind worden. Voor een verzekeraar is het formeel valideren van modellen van cruciaal belang om te achterhalen of modellen ook doen waarvoor ze ontwikkeld zijn ('fit-for-purpose'). Bij onvoldoende validatie bestaat het risico dat het model niet langer fit-for-purpose is, of zelfs dat de verzekeraar geen criteria heeft aan de hand waarvan het model al dan niet als fit-for-purpose geclassificeerd kan worden.

Dit risico geldt in zijn algemeenheid voor alle modellen. Echter voor machine learning modellen, die voor hun kalibratie afhankelijk zijn van de inputdata waarmee ze getraind worden, is periodieke hertraining met nieuw beschikbare data des te wenselijker. Zulke hertraining kan bijdragen aan het 'fit-for-purpose' houden van de modellen, maar het roept tevens vragen op. Via hertraining kan het model aanzienlijk veranderen: wanneer zijn dergelijke wijzigingen dusdanig materieel dat in feite gesproken kan worden van een 'nieuw' model, waarop vervolgens formele validatie moet worden toegepast? Het risico is dat validatie onvoldoende gelijke tred houdt met de ontwikkelingen in het model en de verzekeraar niet langer in-control is.

Hier moet echter wel een afweging worden gemaakt, tussen het in-control blijven over de modellen enerzijds; en de praktische haalbaarheid van hervalidaties anderzijds. Het is van belang dat verzekeraars hierover nadenken en 'grote' en 'kleine' veranderingen in het model verder definiëren. Ook kunnen verzekeraars minimale frequenties van hervalidaties vaststellen.

Aandachtspunt 8 - Validatie

Het vaststellen van een validatieprocedure voor AI-algoritmes/-modellen is van groot belang. Antwoorden op onderstaande vragen kunnen bijdragen bij het opstellen van deze procedure:

- Hoe belangrijk/kritiek is het model qua impact op de klant en stabiliteit van de verzekeraar en is de validatieprocedure toegespitst op het belang van het model?
- Hoe wordt in de validatieprocedure gedifferentieerd tussen verschillende types machine learning technieken, tussen verschillende trainingswijze van modellen (supervised, unsupervised learning) en tussen zelflerende en niet-zelflerende algoritmes?
- Hoe worden 'grote' en 'kleine' veranderingen in het model gedefinieerd? Hoe groot dienen de veranderingen te zijn voordat een formele hervalidatie plaats dient te vinden?
- Welke criteria en situaties worden gebruikt om vast te stellen of het model, de data of aannames zoals gebruikt niet langer passend zouden zijn?
- Wat is de rol van scenario-analyse – waarbij de performance van het model wordt getest bij extreme scenario's (extreme inputdata) die niet in de trainingsdata zijn verwerkt – in de validatieprocedure?
- Hoe wordt de kwaliteit (juistheid, volledigheid, passendheid) van de inputdata meegenomen in de validatie?
- Voor een niet-zelflerend algoritme/model: wat is de frequentie waarmee een validatie tenminste plaatsvindt?
- Voor een zelflerend algoritme/model: hoe kan het trainingsproces als zodanig worden gevalideerd? Wat is de rol voor continue sanity checks en eventueel outputbeperkingen?

Voor zelflerende algoritmes/modellen zijn de vragen weer anders: hier vindt in beginsel een constante update plaats van het model. Dit betekent dat een periodieke validatie weinig tot geen nut heeft. Validatie zou daarom moeten plaatsvinden op het proces waarmee een model continu wordt aangepast. Onderdeel hiervan zijn tevens sanity checks op de uitkomsten van het model, ook gericht op de vraag of de uitkomsten plausibel zijn. Eventueel kunnen ook 'vangrails' worden ingesteld, waarbij voor uitkomsten die buiten een vooraf-gestelde bandbreedte vallen, automatisch verdere controles op en mogelijk aanpassingen van het model zouden volgen.

3.3 AI en de consument

Welke aandachtspunten liggen er rond de zorgplicht bij toepassing van AI?
AI-toepassingen geven verzekeraars een groot aantal mogelijkheden om karakteristieken van consumenten op meer gedetailleerde wijze te analyseren. Eerder is ingegaan op het belang dat verzekeraars technisch in staat zijn de uitkomsten van zulke AI-toepassingen te herleiden en te verklaren. Minstens zo belangrijk is dat AI-toepassingen worden ingezet op een wijze die maatschappelijk acceptabel en uitlegbaar zijn.

3.3.1 Online beslisomgeving

Het is van belang dat AI-toepassingen, gebruikt in beslisomgevingen, consumenten stimuleren (bewust of onbewust) beslissingen te nemen die hun financiële welzijn ten goede komen. De beslisomgevingen betreffen de context waarbinnen consumenten hun beslissingen nemen als het gaat om het al dan niet aanschaffen van een aangeboden product tegen een aangeboden prijs (premie). De verzekeraar heeft op verschillende wijzen invloed op de beslisomgevingen, bijvoorbeeld op het productaanbod of door sturing (nudging), waarbij consumenten indirect richting bepaalde keuzes worden geleid.

Onder de inrichting van de beslisomgeving valt ook gedragsbeprijzing. Gedragsbeprijzing wordt ook wel dynamische beprijzing (dynamic pricing) of beprijzingsoptimalisatie (pricing optimisation) genoemd en betreft het beprijzen van een verzekeringsproduct op basis van het gedrag van consumenten. Dit gedrag staat goeddeels los van het risicoprofiel van de consument en heeft meer betrekking op bijvoorbeeld de kans dat een consument/klant een product afsluit of opzegt (elasticiteit), of de verwachte klantwaarde van een consument voor de verzekeraar. Toenemende beschikbaarheid van (gepersonaliseerde) data en analysetechnieken maken AI-toepassingen in grotere mate inzetbaar voor gedragsbeprijzing. Hiermee ontstaat het risico dat in de beslisomgeving onredelijke variabelen worden gebruikt, of dat de omgeving de consument stuurt richting keuzes die niet per se in het financieel belang van de consument zijn.

Aandachtspunt 9 – Inrichting beslisomgeving

Het is van belang dat AI-toepassingen, gebruikt in beslisomgevingen van de verzekeraar, consumenten stimuleren (bewust of onbewust) beslissingen te nemen die hun financiële welzijn ten goede komen.

Bij het inrichten van beslisomgevingen kunnen de drie volgende overwegingen meegenomen worden:

- Wat mag van een in redelijkheid handelende ('gemiddelde') consument worden verwacht, bijvoorbeeld op het gebied van inspanning om aanbieders te vergelijken of kennis te nemen van product en beprijzing?
- Wat is de impact van een (verkeerde) keuze voor een consument?
- Hoeveel moeite kost het een verzekeraar om de consument te behoeden voor een keuze die niet zijn/haar financieel welzijn bevorderen?

Deze vragen laten tevens het belang zien van beslisomgevingen die dynamisch zijn en zijn toegespitst op de kenmerken van verschillende producten en groepen consumenten. Wat redelijkerwijs mag worden verwacht van de ene groep consumenten – bijvoorbeeld jongeren of hoger opgeleiden – is niet per se redelijk voor andere groepen – bijvoorbeeld ouderen. Ook differentiatie tussen producten verdient in dit licht aandacht: zo is de impact van een verkeerde keuze bij bijvoorbeeld een autoverzekering naar verwachting kleiner dan bij een arbeidsongeschiktheids- of levensverzekering.

Dit alles betekent geenszins dat AI-toepassingen níet ingezet zouden moeten worden in de beslisomgeving. AI-toepassingen kunnen worden ingezet in het productontwikkelingsproces (waarvoor de productontwikkelingsnormen gelden) om het product en de doelgroep beter op elkaar te laten aansluiten, ondersteund door een passende beslisomgeving. Sterker, in sommige gevallen is het mogelijk dat nu of in de toekomst het met behulp van AI-toepassingen dusdanig gemakkelijk wordt om een klant voor bepaalde verkeerde keuzes te behoeden, dat een verzekeraar het risico loopt niet aan haar zorgplicht te voldoen indien ze dergelijke toepassingen **niet** in zou zetten.

3.3.2 Maatschappelijke uitlegbaarheid

Het is van belang dat gebruikte AI-toepassingen en uitkomsten niet alleen technisch, maar ook maatschappelijk uitlegbaar en maatschappelijk acceptabel zijn. Maatschappelijke uitlegbaarheid gaat verder dan het kunnen uitleggen van de techniek of de uitkomsten van AI-modellen. Maatschappelijke uitlegbaarheid raakt aan de vraag of de uitkomsten van verzekeringsmodellen maatschappelijk en ethisch als acceptabel en 'fair' worden gezien.

In dit kader kan worden nagedacht over het type inputparameters dat wordt ingezet. Inputparameters die nu door verzekeraars worden gebruikt zijn statistisch gezien doorgaans zeer significant gecorreleerd met de kans op bijvoorbeeld een bepaalde schade. Ook is er vaak sprake van een intuïtieve oorzaak-gevolg relatie, zodat het voor consumenten helder te begrijpen is waarom een verzekeraar die parameter gebruikt als proxy voor schadekans. Dit verhoogt maatschappelijke acceptatie van de parameter. Ook bij de inzet van AI-toepassingen is het van belang te kijken naar de (intuïtieve) verbanden tussen parameters en het vast te stellen risico.

Tevens is aandacht voor hoe individuen onderverdeeld worden in risicogroepen van belang. Het maken van dergelijke indelingen is een van de kerncompetenties van verzekeraars. Wanneer het risicoprofiel van een persoon op punten afwijkt van dat van de risicogroep waarin die persoon door een verzekeraar is geplaatst, kan zo'n uitkomst zowel technisch uitlegbaar zijn als gebaseerd op intuïtief causale verbanden, maar toch als oneerlijk of maatschappelijk onacceptabel worden beschouwd. Dit geldt in het bijzonder voor gevallen waar momenteel op basis van één kenmerk, bijvoorbeeld beroep, individuen worden uitgesloten van dekking.

AI biedt verzekeraars mogelijkheden om meer verbanden bloot te leggen en daarmee personen langs een groter aantal dimensies in risicogroepen in te delen. Verzekeraars kunnen hierbij de maatschappelijke uitlegbaarheid van uitkomsten van AI-toepassingen waarborgen door uitkomsten van modellen af te zetten tegen maatschappelijke wenselijkheid (bijvoorbeeld solidariteit), door consumenten de

Aandachtspunt 10 – Toetsen toepassing aan maatschappelijke uitlegbaarheid

Bij inzet van AI-toepassingen is het van belang dat de uitkomsten van de toepassing worden afgezet tegen wat maatschappelijk uitlegbaar is:

- In hoeverre is het gebruik van de door de AI-toepassing gevonden en gebruikte verbanden en proxies maatschappelijk gezien fair en uitlegbaar? En hoe is dit getoetst?
- Hoe sterk is de correlatie tussen de gevonden en gebruikte verbanden en proxies en het te verzekeren risico?
- In hoeverre bestaat er een intuïtief verband tussen de gevonden en gebruikte verbanden en proxies enerzijds en de schadekans en risico voor de verzekeraar anderzijds?
- In hoeverre en hoe is het mogelijk voor individuen om afwijkingen van de peer group waarin men is geplaatst te identificeren, aan te tonen en aan te kaarten?
- In hoeverre is de keuze voor verdergaande of minder verdergaande microsegmentatie maatschappelijk uitlegbaar?

mogelijkheid te geven of te helpen hun risico's aantoonbaar te verlagen; en ook door consumenten inzicht te geven in hoe hun risicoprofiel afwijkt van dat van hun *peer group*.

AI kan als zodanig ook helpen de maatschappelijke uitlegbaarheid te vergroten: het geeft verzekeraars de kans om 'macro' risicogroepen te verfijnen tot op (bijna) individueel niveau. Dergelijke microsegmentatie kan ertoe leiden dat voor individuele personen acceptatie- en premiebesluiten beter aansluiten bij hun individuele risicoprofiel. Het geeft individuen ook de kans te laten zien dat hun risicoprofiel mogelijk afwijkt van dat van hun peer group. Een recent praktijkvoorbeeld hiervan betreft de verhogingen van premies of aanscherpingen van voorwaarden voor taxiverzekeringen. Een verfijning van de risicoprofielen van taxichauffeurs aan de hand van inzet van AI (bijvoorbeeld data-analyse van IoT-data) kan leiden tot een beter onderscheid tussen chauffeurs met een laag en een hoog risico. Dit kan niet alleen leiden tot betere voorwaarden voor chauffeurs met een laag risicoprofiel, maar geeft chauffeurs ook de kans om invloed uit te oefenen op de uitkomst door maatregelen te nemen om hun risicoprofiel te verlagen.

Wel heeft verfijndere risicosegmentatie mogelijk een negatief effect op het solidariteitsprincipe. Hier zal in hoofdstuk 4 verder op worden ingegaan.

4. Effect van AI op solidariteit

Afhankelijk van de toepassing kan AI zowel een positieve als negatieve impact hebben op verzekerbaarheid en solidariteit. De verwachte groei in de inzet van AI benadrukt de noodzaak van een breder maatschappelijk debat over solidariteit in de verzekeringssector.

Door toenemend gebruik van AI bestaat de kans dat de solidariteit tussen verschillende risicogroepen in een verzekering onder druk komt te staan.

Dit dilemma is niet nieuw en niet uitsluitend het gevolg van AI-toepassingen: de mate van gewenste solidariteit is een bredere en oudere discussie.

Verregaande toepassing van AI kan dit dilemma wel beïnvloeden: wanneer verzekeraars met behulp van meer data en krachtigere modellen risico-inschattingen op persoonlijker niveau kunnen maken, kan dat ook doorwerken in de premie. De verschillen in premies kunnen groter worden naar mate de verschillen in risico's tussen personen op meer granulair niveau kunnen worden bepaald.

Deze ontwikkeling hoeft echter niet noodzakelijkerwijs te leiden tot verlies aan solidariteit:
Allereerst kunnen AI-toepassingen informatie-asymmetrieën verkleinen, waardoor de verzekerbaarheid van groepen die eerst niet of moeilijker te verzekeren waren in de toekomst zou kunnen verbeteren.

Ook is het niet noodzakelijkerwijs oneerlijk wanneer personen met hogere risico's meer betalen. Dit geldt vooral wanneer de persoon in kwestie aantoonbaar invloed heeft op zijn of haar risicoprofiel en zo ook de premie kan beïnvloeden. Hier kan de inzet van AI de solidariteit zelfs versterken: het vertrouwen dat anderen hun steentje bijdragen om risico's in de risk pool te verlagen kan erdoor toenemen.

Ook bij producten waarbij het moeilijk is voor verzekerden om hun risico's te verlagen, hoeft het inzetten van AI voor beprijzingsdoeleinden niet te leiden tot een wezenlijke afname van solidariteit. Zo is bij een arbeidsongeschiktheidsverzekering nu ook al bekend dat sommige beroepen risicovoller zijn dan andere en wordt op basis hiervan ook geprijsd. AI heeft de potentie verzekeraars in staat te stellen nauwkeuriger te kijken naar de persoon als geheel (en niet alleen naar de parameter 'beroep'), waardoor voor sommige verzekerden de premie relatief lager zou kunnen uitvallen dan onder de huidige beprijzingsmethodes het geval is.

Toch ligt, zeker bij verzekeringen als arbeidsongeschiktheidsverzekering of overlijdensrisicoverzekeringen het risico op de loer dat meer gepersonaliseerde risico-inschattingen – in ieder geval voor een deel van de consumenten – leidt tot hogere premies of uitsluiting. Zeker wanneer risico's niet of deels beïnvloed kunnen worden door de consument, kan dit solidariteit in de sector ondermijnen.

In de discussie over AI en solidariteit is het allereerst vooral de sector die aan zet is: verzekeraars kunnen en moeten afwegen hoe de inzet van AI-toepassingen zich verhoudt tot solidariteit (zie ook aandachtspunt 10). Initiatieven vanuit de sector – zoals de Solidariteitsmonitor, die jaarlijks door het Verbond van Verzekeraars wordt ontwikkeld en als doel heeft ontwikkelingen op het terrein van solidariteit in de gaten te houden – zijn zeer welkom. De AFM en DNB denken hierin ook graag mee en zijn bereid de dialoog verder aan te gaan.

Wel is het van belang dat de maatschappelijke discussie over solidariteit breder wordt gevoerd en zich niet alleen richt op AI-toepassingen: er zal gesproken moeten worden over de gewenste mate van solidariteit in het verzekeringsstelsel als geheel. Het gaat hier om keuzes die de maatschappij als geheel raken en die de samenleving dus als geheel zal moeten maken.

Begrippenlijst

Artificiële intelligentie: Toepassingen die zijn gebaseerd op het analyseren van gevarieerde en grote hoeveelheden data met behulp van technieken als machine learning.

De technieken die worden gebruikt in AI-toepassingen (zoals machine learning) zijn 'intelligent' in de zin dat ze in staat zijn rationeel te optimaliseren: binnen een gegeven taak zijn ze in staat de beste actie te kiezen om een bepaald doel te bereiken, rekening houdend met gestelde criteria.

Naast het gebruik van intelligente technieken is ook data een voorwaarde voor en integraal onderdeel van AI-toepassingen. De ontwikkeling van AI wordt gevoed door de grotere volumes en diversiteit aan beschikbare data. In de verkenning zal het begrip 'big data' worden gebruikt om te verwijzen naar deze toename van inzetbare data.

Machine learning: Een breed onderzoeksveld dat onderdeel is van artificiële intelligentie, waarbij de focus ligt op het verbeteren van de prestatie van een systeem door het systeem te trainen aan de hand van optimalisatie-algoritmes en ingegeven data.

Unsupervised learning: Een vorm van machine learning waar de machine leert van inputdata die niet geclassificeerd, gelabeld of gecategoriseerd is. In plaats van te reageren op feedback, worden overeenkomsten in de gegevens geïdentificeerd en reageert het op basis van de aan- of afwezigheid van dergelijke overeenkomsten in elk nieuw gegeven.

Supervised learning: Een andere benadering binnen machine learning, waar de machine wordt getraind met vooraf beschikbare input en output. Deze data dient als leerbasis voor de machine om een toekomstige taak met soortgelijke inputdata uit te kunnen te voeren.

Reinforcement learning: Een vorm van machine learning geïnspireerd op de menselijke gedragspsychologie, waarbij een machine leert door beloond te worden bij correcte prestaties en gestraft wordt voor verkeerde uitvoer. De machine leert, zonder menselijke tussenkomst, de beloning te maximaliseren en straf te minimaliseren.

Deep learning: Een subset van machine learning, met het verschil dat deep learning op een niet-lineaire manier een probleem oplost in tegenstelling tot machine learning waar een lineair proces wordt doorlopen. Er wordt gesproken van 'deep' learning, omdat het proces door meerdere lagen heen gaat alvorens er een uitkomst wordt gegenereerd.

Deep neural networks: Deep neural networks (DNN) zijn netwerken waarbij de statistische output van de ene laag wordt omgezet in input voor de volgende laag (vandaar de term deep). Of input in een volgende laag wordt gebruikt hangt ook af van drempelwaardes die worden meegegeven in het model.

Clustering: Clusteringtechnieken zijn voorbeelden van unsupervised learning. Deze technieken worden gebruikt om overeenkomsten tussen datapunten (bijvoorbeeld correlaties) te vinden en datapunten vervolgens te groeperen. Dit maakt het bijvoorbeeld bruikbaar om e-mails te clusteren op basis van verbanden of gebruikte woorden in de tekst.

Anomaly detection: Betreft een geautomatiseerd proces dat gegevens die niet thuishoren in een set of patroon identificeert. Verzekeringsmaatschappijen passen deze techniek voornamelijk toe om fraude te detecteren bij claims van klanten. Claims met een afwijkend patroon worden eruit gefilterd en doorgezet naar een medewerker die de claim beoordeelt.

Gradient boosting: Deze techniek wordt toegepast om patronen ('bias') in de 'error term' van andere modellen of predictors te ontdekken en daarmee de error term te verkleinen en zo voorspellingen te verbeteren.

Random forests: Zoals de naam al doet vermoeden, combineert dit type model een groot aantal beslisbomen, om uiteindelijk te komen tot de best mogelijke uitkomst. Dergelijke modellen worden relatief veel gebruikt, deels omdat ze bruikbaar zijn bij relatief beperkte hoeveelheden data. Ook is het tot op zekere hoogte mogelijk om het 'gewicht' van iedere parameter te meten.

Natural Language Processing: Betreft een toepassing, van onder andere artificiële intelligentie, dat zich bezighoudt met de interacties tussen een computer en de mens. Focus ligt op hoe computers moeten worden geprogrammeerd voor het verwerken en analyseren van grote hoeveelheden natuurlijke taalgegevens.

False positives: De uitkomst komt niet overeen met de werkelijkheid. Een voorbeeld van een false positive in claim management is een afgekeurde claim die eigenlijk goedgekeurd had moeten worden op basis van de gebruikte data van de claim.

False negatives: De uitkomst komt niet overeen met de werkelijkheid. Een voorbeeld van een false negative in claim management is een goedgekeurde claim die eigenlijk afgekeurd had moeten worden op basis van de gebruikte data van de claim.

Underfitting: Het algoritme (model) is niet in staat trends in de data te beschrijven. Het betreft hier dan een ondermaatse 'fit' op de gebruikte dataset.

Overfitting: Het model is te specifiek getraind op een bepaalde dataset waardoor het model weliswaar een goede fit is voor die dataset, maar niet voor andere datapunten.

Spurieuze verbanden: Een situatie waar twee punten met elkaar kunnen correleren, terwijl ze afhangen van een ander punt. Hierdoor ontstaat een zogenaamd schijnverband tussen twee punten, er is namelijk geen duidelijk oorzaak-gevolg (causale) relatie.

Bias: Vertekening van uitkomsten door systematische of toevallige fouten. Mensen kunnen ook bias vertonen, wat het leidt tot een irrationele, foutieve gedachtegang. Dit betreft een bevooroordeling die zowel onbewust als bewust kan voorkomen.

Explainer model: Maakt inzichtelijk welke keuzes een model heeft gemaakt om te komen tot de uitkomst die het presenteert.

Partial dependence: Visualiseert hoe elke variabele of voorspeller de voorspellingen van het model beïnvloedt.

Black box: Een complex systeem of algoritme waarvan de interne werking verborgen is of niet gemakkelijk te begrijpen is.

Proxy: Een variabele wat een afgeleide of verklaring van een andere, niet meegenomen, variabele kan zijn.

Referenties

Actuaries Institute (2016), 'The impact of Big Data on the Future of Insurance', November 2016

Balasubramanian, R., Libarikian, A., McElhaney, D. (2018), 'Insurance 2030 – The impact of AI on the future of insurance', McKinsey

Bank of America & Merrill Lynch (2018) Robotics and artificial intelligence (AI) worldwide market size estimates, based on 2018 to 2030 forecasts, by segment (in billion U.S. dollars)

Cai, E. (2014), 'Machine Learning Lesson of the Day – Overfitting and Underfitting', StatBlogs

College van de Rechten van de Mens (2018) 'Wat is discriminatie', beschikbaar via https://www.mensenrechten.nl/nl/discriminatie-uitgelegd

The Geneva Association (2018) 'Big Data and Insurance: Implications for Innovation, Competition and Privacy', March 2018

Gan, G. (2013), 'Application of data clustering and machine learning in variable annuity valuation', Insurance: Mathematics and Economics, Vol. 53 (2013), pp. 795-801

Grand View Research (2017), 'Artificial Intelligence Market Analysis By Solution (Hardware, Software, Services), By Technology (Deep Learning, Machine Learning, Natural Language Processing, Machine Vision), By End-use, By Region, and Segment Forecasts, 2018 – 2025'.

Helveston (2016) 'Consumer Protection in the Age of Big Data', Washington University Law Review

Joint Committee of the European Supervisory Authorities (2018), 'Joint Committee Final Report on Big Data', JC/2018/04, 15 March 2018

Kearns, M. (2017), 'Fair Algorithms for Machine Learning', 18th ACM Conference on Economics and Computation 26-30 juni 2017

Mullainathan, S., Spiess, J. (2017), 'Machine learning: an applied econometric approach', Journal of Economic Perspectives Vol. 31-2, Spring 2017, pp. 87-106

Nian et al. (2016), 'Auto insurance fraud detection using unsupervised spectral ranking for anomaly', The Journal of Finance and Data Science

O'Neil, Helfand, Kochenburger (2017) 'Big Data: Looking Under the Hood and Into the World', Pullman & Comey Attorneys LLC, 18 september 2017

Orbis Research (2017), 'Global Artificial Intelligence Market Analysis and Forecast 2022 by Size, Share and Growth Rate'.

Roy, R. and Thomas George, K. (2017), 'Detecting insurance claims fraud using machine learning techniques', presented at 2017 International Conference on Circuit, Power and Computing Technologies.

Seely, S. (2018), 'Eight use cases of for machine learning in insurance', Microsoft Azure

Tata Consulting Services (2017), 'Hoe Kunstmatige Intelligentie de financiële dienstverlening verandert'

TechEmergence (2017) Share of companies investing in artificial intelligence (AI) worldwide, by industry, as of 2016, en Tata Consultancy Services

Tractica (2017), 'Artificial Intelligence Software Market to Reach $89.8 Billion in Annual Worldwide Revenue by 2025'.

Verbond van Verzekeraars (2018), 'Voertuigdata: de klant aan het stuur'

Verbond van Verzekeraars (2016) 'Grip op Data'

Vetzo, Gerards, Nehmelman (2018) 'Algoritmes en grondrechten', Boom Uitgevers, Den Haag

Wang, Y, Xu, W. (2019), 'Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud' Decision Support Systems, Vol. 105 (Jan 2018), pp. 87-95

Zhang et al (2017), 'Understanding deep learning techniques rethinking generalization', Cornell University Paper