



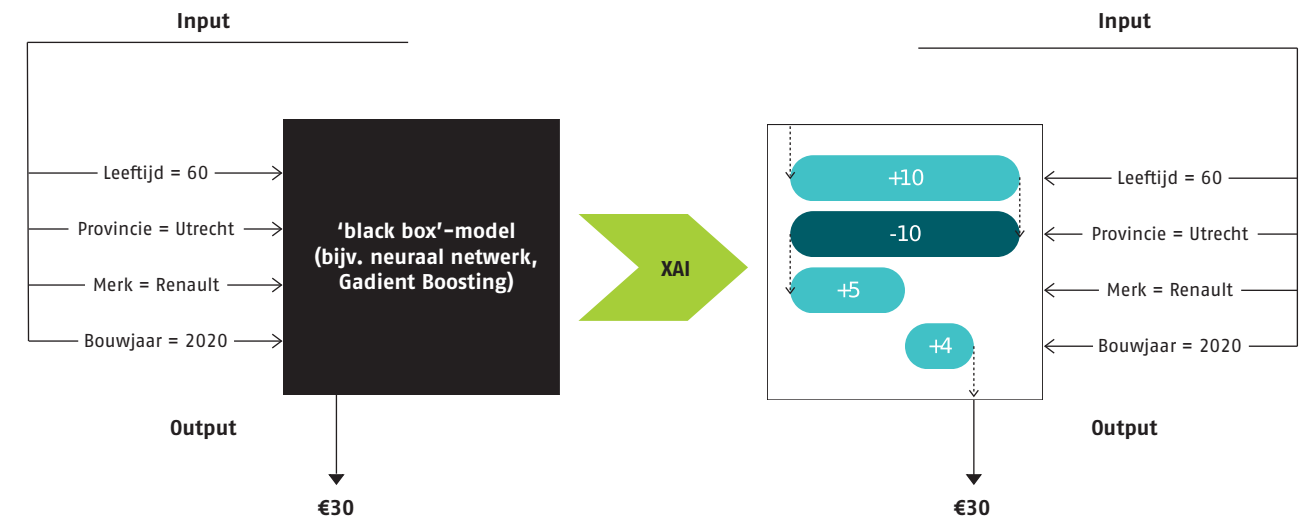
De 'black box' uitgepakt: de mogelijke rol van Explainable Artificial Intelligence voor de verzekeringswereld

Kunstmatige intelligentie (AI) is een onmisbare trend geworden, maar deze snelle ontwikkeling brengt ook diverse uitdagingen met zich mee. Om AI op een waardevolle manier in te zetten voor de verzekeringswereld is het belangrijk dat bepaalde uitdagingen worden overwonnen, zoals het 'black box'-karakter van de techniek. Dit artikel zal dieper ingaan op hoe Explainable Artificial Intelligence (XAI)-technieken de toepassing van AI binnen de verzekeringsbranche kunnen verbeteren en uitdagingen zoals 'black box'-algoritmes mogelijk kunnen oplossen.

Consumenten en toezichthouders eisen steeds vaker dat de besluitvormingsprocessen van AI transparant, duidelijk en rechtvaardig zijn. Daarnaast is de eerlijkheid van een model een belangrijke parameter die inzichtelijk moet zijn tijdens de ontwikkeling en implementatie van zo'n model. Met XAI kan men duidelijke en begrijpelijke inzichten bieden in de besluitvormingsprocessen van complexe AI-algoritmes, en opent daarmee de 'black box'-structuur van deze AI-modellen. Het uiteindelijke doel is om AI-systemen te creëren die gemakkelijk door mensen begrepen en geïnterpreteerd kunnen worden. Binnen de wereld van AI moet er op dit moment een keuze worden gemaakt tussen uitlegbaarheid of een hoog prestatieniveau. Goed uitlegbare modellen, zoals lineaire/logistische regressies, zijn vaak niet in staat om een hoge accuraatheid te bereiken. Aan de andere kant zijn neurale netwerken juist heel accuraat, maar daar is het door een grote hoeveelheid aan (intergeconnecteerde) lagen lastig om inzichtelijk te krijgen hoe de output bepaald is.

XAI is een verzameling van technieken die, in vergelijking tot neurale netwerken, uitlegbare modellen produceren én daarbij een hoog prestatieniveau behouden. Enkele voorbeelden van zulke XAI-technieken zijn LIME (Local Interpretable Model-agnostic Explanations) en SHAP (SHapley Additive exPlanations). LIME richt zich op het uitlegbaar maken van individuele voorspellingen. Het onthult hoe kleine aanpassingen aan specifieke datapunten de voorspelling van een model kunnen beïnvloeden. Dit helpt gebruikers om te begrijpen hoe een model zou kunnen presteren onder verschillende omstandigheden rondom een specifiek voorbeeld. SHAP biedt, naast het inzicht in hoe individuele voorspellingen tot stand komen, ook inzicht in de globale impact van kenmerken binnen een model. Dat maakt dat SHAP een completere XAI-techniek is dan LIME. Hoe SHAP inzicht geeft in een individuele voorspelling, is zichtbaar in het figuur. Stel dat we een 'black-box'-premiemodel hebben. Om tot een premie voor één individu te komen, wordt het premiemodel met een aantal kenmerken van een individu gevoed. Een standaard black-box model zal een premie geven, zonder dat duidelijk is hoe het model tot deze output is gekomen (linkerkant). Met behulp van SHAP kan het plaatje aan de rechterkant worden geconstrueerd. Hierin kan precies worden nagegaan welke kenmerken van een individu én in hoeverre deze bijdragen aan de output van het model. Zo zorgt in dit voorbeeld het hebben van een leeftijd van 60 voor een effect van +10 op de premie (ten opzichte van de gemiddelde output van het model).

J. Reil MSc (links) is een PhD kandidaat bij de Bern Business School, maar schreef dit artikel in dienst van Triple A – Risk Finance, waar ook consultant dr. T. Jacobs (midden) en senior consultant A. Pison MSc werken.



Figuur 1: een standaard black-box-voorspelling (links) versus een uitlegbare XAI-voorspelling (rechts)

Een andere interessante techniek binnen het XAI-domein is de zogeheten 'Explainable Boosting Machine' (EBM). EBMs zijn ontworpen om de hoge nauwkeurigheid van traditionele *boosting machines* te behouden, terwijl ze tegelijkertijd de uitlegbaarheid verbeteren. EBMs gebruiken een additief model dat bestaat uit een reeks eenvoudige, interpreteerbare componenten (beslisbomen). Dit maakt het mogelijk om de bijdrage van elke component afzonderlijk te begrijpen. Een EBM is in feite een 'glass box'-model en heeft dus geen additionele technieken zoals SHAP of LIME nodig om uitlegbaar te zijn. Door deze eigenschappen zijn EBMs interessant voor verzekeringsmaatschappijen die willen begrijpen hoe hun AI-modellen tot bepaalde beslissingen komen, zonder dat ze moeten afzien van de nauwkeurigheid van hun voorspellingen.

Met de inwerkingtreding van de Europese AI Act, die sinds 1 augustus 2024 van kracht is, is de rol van XAI steeds belangrijker geworden. De AI Act stelt namelijk eisen aan de ontwikkeling en het gebruik van AI-systemen om ervoor te zorgen dat AI veilig, betrouwbaar en transparant is. De AI-toepassingen worden geclassificeerd op basis van bepaalde risico's en de AI Act stelt daarbij verschillende eisen afhankelijk van het risiconiveau. Voor hoog-risico AI-systemen, zoals die gebruikt worden in de financiële dienstverlening, zijn duidelijke eisen gesteld om de risico's voor fundamentele rechten te minimaliseren. XAI helpt hierbij door de werking van deze systemen te verduidelijken, zodat fouten sneller kunnen worden opgespoord en gecorrigeerd. Daarnaast moeten ontwikkelaars van AI-systemen de risico's voor fundamentele rechten in kaart brengen en klanten actief informeren over de inzet van hoog-risico AI-systemen. Dit zorgt ervoor dat klanten bewust zijn van wanneer AI wordt gebruikt in beslissingen die hen direct raken, zoals bij financiële beoordelingen. Door de besluitvormingsprocessen transparanter te maken, kunnen verzekeringsmaatschappijen klanten beter informeren over hoe bepaalde beslissingen (bijvoorbeeld claims of premieberekeningen) tot stand komen. Dit bevordert niet alleen de klanttevredenheid, maar helpt ook bij het naleven van regelgeving, aangezien de verantwoording van beslissingen hiermee vereenvoudigd wordt. Bovendien stelt XAI verzekeringsbedrijven in staat om ethische overwegingen systematisch te integreren in hun AI-modellen, zodat deze niet alleen efficiënt, maar ook rechtvaardig zijn.

De integratie van XAI in de huidige datacultuur kan dus zowel uitkomsten bieden voor de klant, als voor de modeldocumentatie én het contact met de toezichthouder. In onze projecten bij verschillende opdrachtgevers zien we wel dat er nog een slag om de arm gehouden

moet worden, want er kunnen ook enkele nadelen verbonden zijn aan het toepassen van XAI. Er moet altijd een afweging gemaakt worden tussen de uitlegbaarheid en de prestaties van het model, zodat de uitkomsten uit het model de klant op een zo'n goed mogelijke manier kunnen faciliteren. De implementatie van XAI, maar ook het onderhouden van de systemen, kan daarnaast ook voor de nodige investeringen zorgen (zowel in mankracht als financieel). De efficiëntie waarmee modellen nu ontwikkeld en geïntegreerd zijn binnen een verzekeraar kan minder worden door XAI, omdat het openen van de 'black box' gepaard gaat met een uitbreiding van de modellen. Tot slot is het geen *plug-and-play* oplossing, maar zal er per model gekeken moeten worden naar de mogelijke kansen en risico's bij het gebruik van XAI.

De integratie van XAI in de verzekeringsindustrie biedt een veelbelovende oplossing voor de complexe uitdagingen waar deze sector momenteel mee geconfronteerd wordt. Door het verbeteren van de transparantie en begrijpelijkheid van AI-systemen, helpt XAI niet alleen om de vertrouwensrelatie met klanten te versterken, maar ook om de operationele en ethische uitvoerbaarheid van AI-toepassingen binnen de sector te waarborgen. ■