



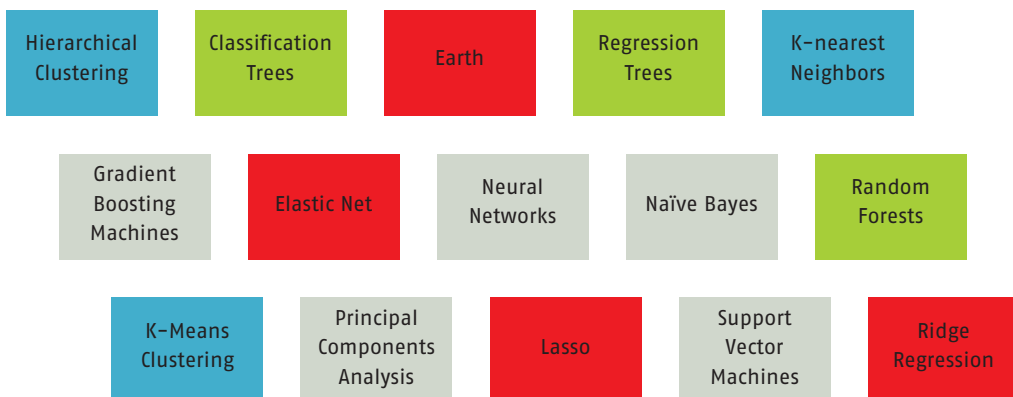
# Machine Learning: een versnellende evolutie



Velen binnen verzekeringsland hebben de termen AI (Artificial Intelligence) en ML (Machine Learning) inmiddels al wel voorbij horen komen de afgelopen jaren. Maar in hoeverre worden AI en ML nu eigenlijk toegepast? Zijn het inmiddels niet gehypte modewoorden geworden waar iedereen het over heeft, maar die toch lastig toepasbaar blijken?

## Machine Learning

Machine Learning is een verzamelnaam voor een aantal technieken waarbij met minimale interventie van menselijke experts beslissingen kunnen worden genomen. Deze technieken kunnen worden ingezet om op een meer geautomatiseerde wijze patronen in historische data te vinden.



### 1. Een greep uit Machine Learning methoden

Machine Learning (ML) technieken worden vaak in twee hoofdgroepen onderverdeeld: supervised en unsupervised. De supervised methoden worden gebruikt om een voorspelling te maken. De unsupervised methoden worden gebruikt om data te groeperen. Voorbeelden van ML methoden zijn Tree-based methoden (Classification trees, Regression trees en Random forests), Penalised regression methoden (Lasso, Ridge, Elastic Net, Earth – en GLM is in wezen een simpelere versie van deze methode) en Clustering methoden (K-Means Clustering, K-Nearest Neighbors en Hierarchical Clustering). Het is afhankelijk van de data en het doel van de modellen of de ML modellen beter zijn dan traditionele modellen zoals de GLM.

J.A.H. Bruins MSc AAG (links) is Manager; drs. P.W.F. Hoogveld (midden) is Senior Director; en J.A. Wilbers MSc AAG is Senior Consultant; allen werken ze bij Willis Towers Watson.

Dit artikel is op persoonlijke titel geschreven.



**Penalised regression**

Door een aantal penalised regression methoden als volgt in formule samen te vatten is goed te zien dat de GLM in wijze een simpele versie van de penalised regression methoden is:

$$f(x) = g^{-1}(x \cdot \beta)$$

waar  $\beta$  verkregen wordt door het volgende te minimaliseren:

$$L(\beta | x, y) + \lambda_1 \sum_i |\beta_i| + \lambda_2 \sum_i \beta_i^2$$

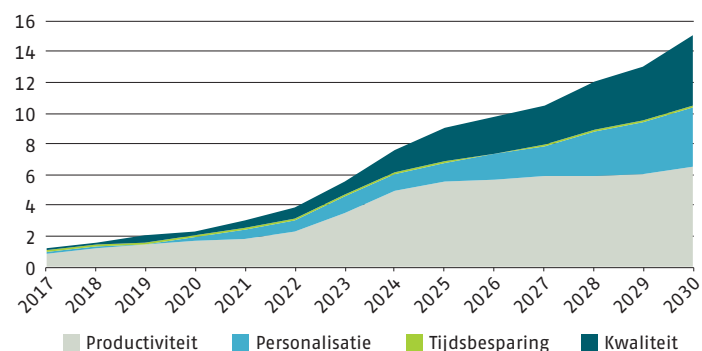
Hierbij is het zwart omlijnde gedeelte herkenbaar als een standaard GLM. De blauwe toevoeging is de Lasso-methode (least absolute shrinkage and selection operator), de rode is de Ridge methode en de combinatie (groen omrand) is de Elastic Net methode.

### Artificial Intelligence

Artificial Intelligence is de overkoepelende wetenschap die zich bezighoudt met het creëren van machines die een vorm van intelligentie vertonen. Machine Learning is hier onderdeel van en helpt mee aan het creëren van de kunstmatige intelligentie.

### WAT ZIJN DE VOORUITZICHTEN?

De verwachting is dat AI de komende jaren een enorme groei zal doormaken en in 2030 \$15 biljoen zal bijdragen aan de wereldwijde economie. Hiermee zal AI zorgen voor veel verandering en is het een van de grootste commerciële kansen van dit moment.



**Figuur 1.** Productiviteit | Personalisatie | Tijdsbesparing | Kwaliteit

Deze verandering zal veel van organisaties vragen en zeker ook van verzekeraars. Digitale transformatie vereist een cultuuromslag die verder gaat dan technologische updates of het herontwerpen van processen om echt van de verwachte voordelen te profiteren. Hoewel, zoals reeds in recente edities van De Actuaris is benoemd, het werk van academische actuarissen niet gemakkelijk te automatiseren is, zal de actuaris zich wel moeten aanpassen aan de nieuwe mogelijkheden en technieken.

Uit onderzoek is gebleken dat er veel ambitie is bij verzekeraars met betrekking tot het gebruik van ML en AI, maar dat men deze ambitie de afgelopen jaren over het algemeen niet heeft kunnen waarmaken (zie tabel 2). Op gebieden als het bouwen van risicomodellen is het gebruik van AI en ML in twee jaar verdubbeld, maar was de ambitie vooraf nog groter. Toch is de verwachting voor heel 2021 dat het gebruik gemiddeld drie keer zo hoog zal zijn als in 2019. De ambitie en verwachting van verzekeraars om de ML- en AI-technieken verder in te zetten blijven dus groeien.

	Actual for 2017	Expected for 2019 (in 2017)	Actual for 2019	Expected for 2021
Build risk models for better decision making	13%	44%	26%	60%
Reduce time spent by humans	11%	49%	22%	60%
Better understand risk drivers	21%	44%	20%	56%
Identify cases that pose higher risk	11%	46%	14%	50%
Augment human-performed underwriting	7%	37%	7%	47%
Identify patterns of fraudulent claims	9%	39%	17%	47%
Identify bottlenecks in claim processes/Process claims more efficiently	3%	30%	7%	43%

2: Gebieden waar verzekeraars de meeste AI- en ML-toepassingen verwachten

### WAT ZIJN DE UITDAGINGEN?

Wanneer we kijken naar de historie van de toepassing van wiskundige technieken, dan zien we een golfbeweging die ontstaat door de beschikbaarheid van (schone) data, IT-capaciteit, kennis, wetgeving en algehele acceptatie. Kort gezegd wakkert de markt de drang naar meer data aan, wat dan zorgt voor investeringen in nieuwe analytische methoden, wat vervolgens weer leidt tot nieuwe inzichten, nieuwe marktbehoeften en dan ook weer nieuwe data. De beschikbare data en de beschikbare analytische methoden zijn eigenlijk nooit perfect in evenwicht.

Tegenwoordig gebruiken verzekeraars veel externe data zoals credit scores, geografische data en social media data. Het gebruik van deze externe data werd mogelijk door meer IT-capaciteit. Elk nieuw type databron moet echter geanalyseerd worden met 'slimmere' technieken en vereist meer capaciteit en kunde. Bijvoorbeeld, modellen analyseren op basis van tien variabelen is een stuk minder tijdrovend dan een model op basis van honderden variabelen. De uitdaging hierbij voor verzekeraars zit niet in de technieken zelf, maar in het vermogen van de verzekeraar om de technieken te benutten.

Daarnaast heeft door het gebruik van al deze data de wetgeving zich aangepast en is bijvoorbeeld de wet AVG (GDPR) ingevoerd waar een verzekeraar rekening mee moet houden.

Overigens wordt ook in de zoektocht naar meer en ook betere data ML toegepast. Zo zien wij dat methoden zoals Text Processing en Text Mining kunnen helpen om al beschikbare – ook niet kwantitatieve – data bruikbaar te maken.

### Waarom zullen AI en ML technieken steeds belangrijker worden?

De ontwikkeling van nieuwe technieken zal zich naar verwachting golvend omhoog voortbewegen, waarbij meer data leidt tot nieuwe inzichten uit voortschrijdende analytische methoden, die op hun beurt weer leiden tot innovaties, wellicht nieuwe wetgeving en ook aanvullende databehoeftes.

AI en ML zullen een steeds prominentere plek krijgen in het verzekeringsbedrijf gezien het momentum dat er is om hierin te investeren en gelet op de te verwachten toegevoegde waarde die deze investeringen zullen opleveren. Naast het feit dat ML-technieken kunnen bijdragen aan accuratere modellen – en daarmee bijvoorbeeld de risico's beter kunnen worden geprijsd of fraude sneller kan worden gedetecteerd –, zullen zij ook de efficiency en snelheid van werken positief beïnvloeden. Dit zal ook kostenreducties met zich meebrengen.

ML en AI zullen in de nabije toekomst toegepast worden door de gehele verzekeringsketen heen. Van marketing tot acquisitie, klantbehoud, acceptatie, pricing en claim management.

**GBM als toevoeging aan de GLM**

Binnen Pricing moet duidelijk zijn hoe de premie is opgebouwd. Hiermee wordt een grote restrictie gelegd op de structuur van de uitkomst (uitlegbaarheid) en daarmee welke techniek er het beste kan worden toegepast. Wat wij in de markt zien gebeuren, is dat er combinaties van modellen worden gebruikt. Zo zien wij bijvoorbeeld dat bij verzekeraars de toepassing van een GBM (Gradient Boosting Machine) op een GLM voor een beter resultaat zorgt, door effecten te ontdekken die bij het schatten van de stand-alone GLM niet duidelijk zichtbaar waren. Zo wordt de GLM-parameterisering verbeterd, waarbij de uitlegbaarheid wel behouden kan blijven.

### AI EN ML IN DE NEDERLANDSE VERZEKERINGSMARKT

De Nederlandse verzekeringsmarkt liep nooit echt voorop waar het gaat om technologisch vooruitstrevende innovaties of verzekeraars die met revolutionaire nieuwe technieken of producten zich echt onderscheiden.

**Topic Modelling voor grote schades**

Door de toepassing van Topic Modelling – een techniek om in grote stukken tekst bepaalde onderwerpen te herkennen – op schade- en medische rapporten kan informatie hieruit direct gebruikt worden om potentieel grote schades ('jumper claims') in een vroeg stadium te identificeren, zodat deze prioriteit kunnen krijgen bij schadebehandelaars en ook de case reserves hiervoor al in een vroeger stadium aangepast kunnen worden.

Door toenemende concurrentie, aanbod van buitenlandse niche-aanbieders en ontwikkeling van nieuwe producten is de laatste jaren – met name door de mogelijkheden die nieuwe technologie biedt – een inhaalslag aan de gang. Dat dit minder snel gaat dan een aantal jaar geleden werd voorspeld, doet daar niets aan af. Er is nu brede acceptatie dat gebruik van technologie en nieuwe analytische methoden noodzakelijk is om bij te blijven. Samen met automatiseren zal het ontginnen van nieuwe databronnen en analytische technieken voorlopig de digitale agenda domineren. ■

### Bronnen

<https://www.willistowerswatson.com/en-US/Insights/2020/02/machine-learning-today-and-tomorrow>

<https://www.willistowerswatson.com/en-US/Insights/2020/06/using-ai-to-augment-a-digital-operating-model-in-the-post-pandemic-world>

# Actuarieel Instituut

- **Opleidingen**  
complete en compacte deeltijd beroepsopleidingen
- **Trainingen & Leergangen**  
actuele en vakinhoudelijke diepgang
- **In company**  
beroepsopleidingen en trainingen op maat

Geïnteresseerd? ga naar [www.ag-ai.nl/actuarieelinstituut](http://www.ag-ai.nl/actuarieelinstituut)