

Praktische handvatten om met predictive analytics aan de slag te gaan

Op basis van de ervaringen van tien organisaties beschrijft Suzanne Deering op welke wijze organisaties predictive analytics kunnen inzetten en wat de do's en don'ts zijn bij de toepassing ervan.

In de technologie die klant- en burgergedrag kan voorspellen is reeds beschikbaar. Evenals de benodigde data. Predictive analytics biedt organisaties daarmee geheel nieuwe kansen om beter te presteren. Maar hoe introduceer je predictive analytics binnen een organisatie? In dit artikel beschrijft de auteur de ervaringen met predictive analytics van tien middelgrote tot grote organisaties in zowel de publieke als private sector.

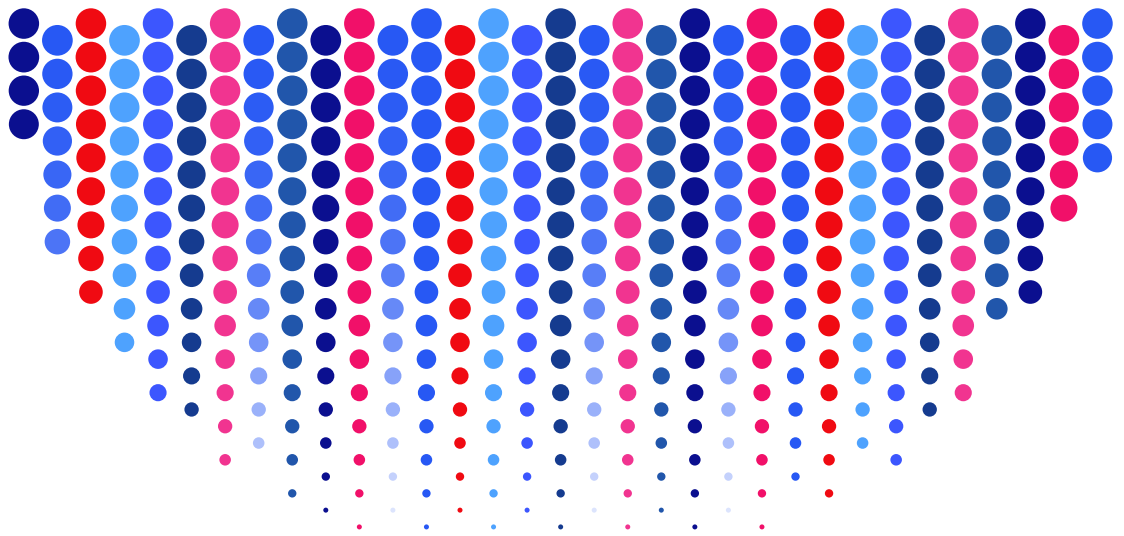
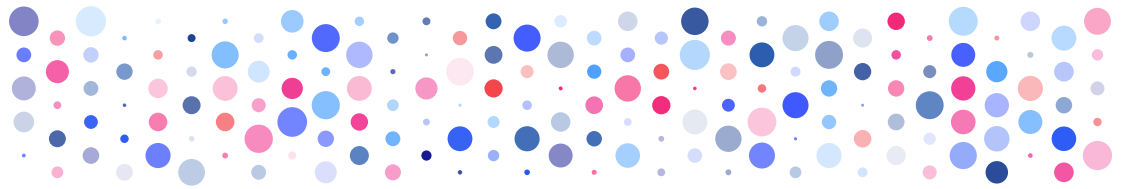
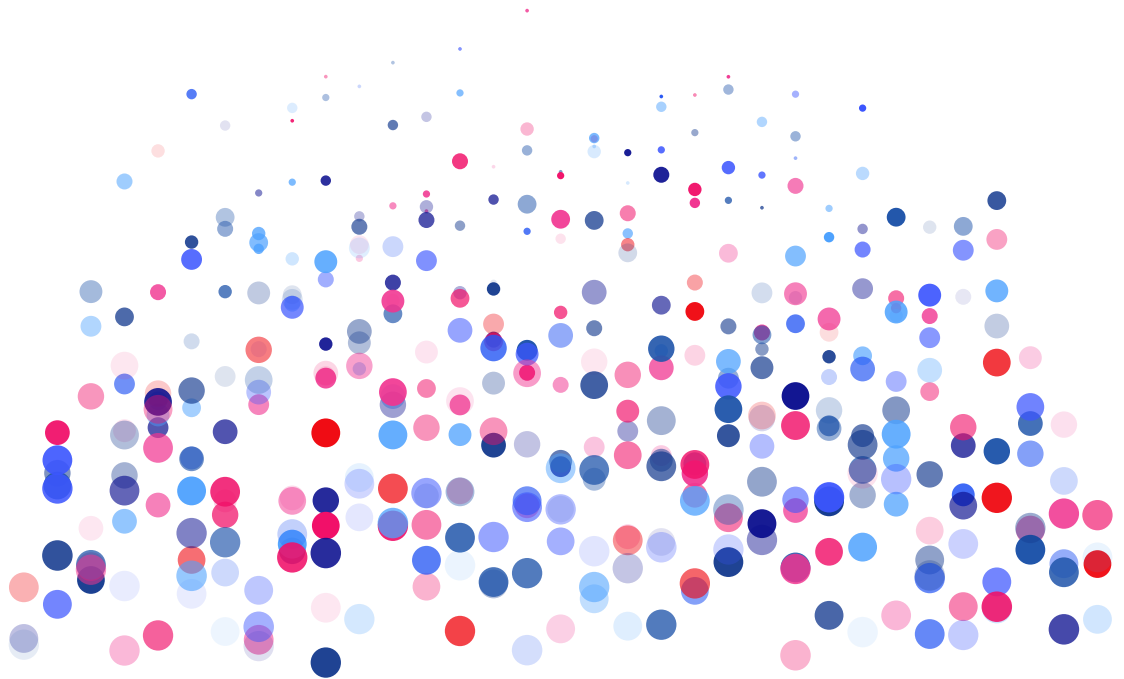
Hoe pak je de introductie van predictive analytics aan?

Grofweg bestaan er twee verschillende aanpakken voor de introductie van predictive analytics: vraaggestuurd en aanbodgestuurd. De vraaggestuurde aanpak stelt de businessvraag centraal; het voorspellingsmodel geeft antwoord op een bestaande businessvraag. De aanbodgestuurde aanpak stelt de beschikbare data centraal; biedt de beschikbare data mogelijkheden om voorspellingen te doen die de organisatie verder helpen?

De bevroegde organisaties adviseren een vraaggestuurde aanpak, tenzij innovatie het kloppende hart

van de organisatie is. Met deze aanpak kan je klein beginnen, ervaring opdoen en een leerproces in gang zetten. De datapreparatie (oftewel het zorgen dat de data compleet, op orde en klaar voor gebruik is) neemt in totaal tachtig procent van de tijd in beslag om een voorspellingsmodel te bouwen. Bij deze aanpak prepareer je alleen de data die relevant is voor de concrete businessvraag. Zo kan je in korte tijd en zonder grote investeringen de toegevoegde waarde van predictive analytics aantonen. Dit zorgt weer voor commitment en ambassadeurschap vanuit het management. En voor budget voor volgende predictive analytics-initiatieven.

Natuurlijk staat het management niet te springen om grote investeringen waar een groot risico aan hangt. De aanbodgestuurde aanpak is om die reden vaak minder aantrekkelijk. Het kost veel geld om alle datasets in kaart te brengen en te prepareren. IT neemt het in kaart brengen van de data vaak voor haar rekening en de business wordt daar veelal niet bij betrokken. Bij de aanbodgestuurde aanpak waarschuwen organisaties dan ook voor onjuiste interpretaties van data en incorrecte analyses wegens het ontbreken van relevante domeinkennis.



Valkuilen omtrent het bouwen van een voorspellingsmodel

Hieronder staan de belangrijkste leerpunten beschreven ten aanzien van de bouw van een voorspellingsmodel:

Verklaarbaarheid

Voorspellingen moeten te verklaren zijn: hoe heeft de data tot de voorspelling geleid? Wanneer het model bijvoorbeeld een grote kans op wanbetaling voorspelt van bepaalde klanten van een bank en hen vervolgens de optie ontnemt om te betalen via acceptgiro, dan dient dit goed te worden onderbouwd. De reputatie van de organisatie kan anders grote schade oplopen. Dit risico zorgt er ook voor dat de kans op acceptatie van het voorspellingsmodel door de business klein is wanneer voorspellingen niet verklaarbaar zijn. Een tip voor de verklaarbaarheid van het model: gebruik lineaire modellen en beslisbomen. De voorspellingen van deze modellen zijn altijd te herleiden tot de data.

Restricties

Stel restricties vast voor het voorspellingsmodel: de grenswaarden waarbinnen het model voorspellingen doet. Een inspectiebureau dat voor dit artikel werd bevestigd, deed dat niet voor een model dat fraudegevallen identificeert. Vooraf was geen bedrag vastgesteld vanaf wanneer het wenselijk zou zijn om fraudegevallen in beeld te krijgen. Zo werden er ook fraudegevallen geïdentificeerd die uiteindelijk meer kosten dan opleverden.

Consequenties van de voorspellingen

Bedenk vooraf of de organisatie in staat is om de uitkomsten te benutten. Het voorspellingsmodel van het reeds besproken inspectiebureau leverde een groter aantal uit te voeren inspecties op. De inspecteurs kregen extra werk en al snel bleek dat het bureau te weinig mankracht had om alle 'nieuwe' inspecties uit te voeren. Stel jezelf vooraf de vraag of bepaalde inzichten wel wenselijk zijn en of je in staat bent om aan bijkomende randvoorwaarden te voldoen.

Presentatie van de voorspellingen

Een beeld zegt méér dan duizend woorden. Hoe voorspellingen gepresenteerd of gevisualiseerd worden heeft een grote impact op de interpretatie ervan. Daarnaast heeft de business vaak specifieke wensen en verwachtingen. Om aan die wensen en verwachtingen te kunnen voldoen, moeten deze tijdens de bouw van het voorspellingsmodel meegenomen worden.

Transparantie

Wees transparant richting de personen op wie het voorspellingsmodel effect heeft. ING heeft aan den lijve ondervonden wat er gebeurt als je dat niet doet. In 2014 haalde de bank de publiciteit met de wens om klantgegevens door te verkopen aan commerciële organisaties. ING kon op basis van de pinhistorie van klanten, zeer gerichte marketing voeren. Had een ING-klant het afgelopen halfjaar veel boeken gekocht, dan was dat enorm waardevolle informatie voor het voorspellingsmodel van bijvoorbeeld Bol.com. ING wilde het doorspelen van deze informatie

Voor veel mensen is het nog een stap te ver om het lot van gevangenen in de handen van 'de computer' te leggen

testen met een kleine groep klanten. Omdat de bank hierover niet transparant was, vreesden klanten voor hun privacy. Transparantie had ING veel negatieve publiciteit kunnen besparen.

Het predictive analytics-team

De correcte interpretatie van data en voorspellingen zijn alleen mogelijk wanneer het predictive analytics-team de juiste domeinkennis heeft. Daarnaast moet de business nauw betrokken zijn bij de bouw van het model, om het aan het management en klanten te kunnen verkopen. Een nauwe samenwerking tussen business en IT is daarom vereist. En om dat te bereiken, moeten organisatiestructuren in de meeste gevallen worden aangepast.

Net als bij het samenstellen van ieder ander team breng je bij het samenstellen van een predictive analytics-team vooraf de benodigde kennis en kunde in kaart en zoek je aan de hand daarvan de juiste teamleden bij elkaar. Vergeet de vertegenwoordiging van juridische expertise niet. Wetgeving rondom het gebruik van persoonsgegevens is voortdurend in beweging en zonder deze expertise is de kans op onrechtmatig handelen groot. Maak ook de afweging welke expertise je in huis wilt hebben en welke expertise je wilt inhuren. Het is vaak niet wenselijk (en mogelijk) om domeinkennis in te huren.

De ethische kant van predictive analytics

Mogelijke gevolgen van predictive analytics voor de organisatiecultuur mogen niet onderschat worden. Wanneer een organisatie met voorspellingsmodellen gaat werken, ervaren medewerkers dit wellicht alsof zij minder invloed krijgen en technologie hun werk overneemt. Dit kan een gevoel van angst oproepen. Zoals het geval bij het eerder besproken inspectiebureau. Voordat het inspectiebureau een voorspellingsmodel in gebruik nam, voerden de inspecteurs

inspecties uit op basis van een onderbuikgevoel. Met de komst van het voorspellingsmodel waren zij niet meer degenen die besloten welke inspecties werden uitgevoerd. Dat had invloed op de eigenwaarde en het 'werkgelek' van de inspecteurs.

Het belang van de verklaarbaarheid van een voorspellingsmodel is al kort aan bod gekomen. Dat belang is des te groter wanneer de voorspellingen betrekking hebben op het lot van mensen. In Amerika is reeds geëxperimenteerd met voorspellingsmodellen die het werk van rechters overnemen. Zo werd er met een voorspellingsmodel bepaald of een gevangene met verlof mocht of niet – afgeleid van de kans op een terugval. Dit heeft tot veel commotie onder de bevolking geleid. Voor veel mensen is het nog een stap te ver om het lot van gevangenen in de handen van 'de computer' te leggen.

De impact van voorspellingsmodellen kan groot zijn - op organisaties, klanten, burgers en de gehele samenleving. Dat vraagt van de gebruikers van deze modellen om goed naar de ethische aspecten te kijken. De consequenties en risico's van het voorspellingsmodel dienen vooraf goed te worden afgewogen. Mijn uitgangspunt is: hoe groter het effect op de mens, hoe groter de invloed van de mens moet zijn op het uiteindelijke besluit. Ga je aan de slag met predictive analytics, bepaal dan eerst wat je ethische standpunt is!

Suzanne Deering is consultant bij Verdonck, Klooster & Associates.

