

Leren van Positieve Afwijkingen

De exponentiële groei van gedigitaliseerde data en rekenkracht zal een steeds grotere invloed hebben op onze maatschappij. Alleen organisaties die slim gebruik maken van de stortvloed aan gegevens over het gebruik van producten en diensten zullen overleven. Er is een nijpend tekort aan “data scientists” – mensen die waarde uit ruwe gegevens kunnen halen door gebruik te maken van data/process-mining technieken, slimme visualisaties, geavanceerde algoritmen en moderne grootschalige IT infrastructuren.

Tegelijkertijd leven we in een “compliance maatschappij” waar auditors organisaties controleren en afwijkingen als ongewenst worden gezien. Dankzij de toegenomen hoeveelheid data en rekenkracht kunnen we het gedrag van individuen en organisaties steeds nauwkeuriger volgen en daardoor steeds beter afwijkingen detecteren. Natuurlijk is het goed dat we dankzij “data science” technieken fraude op kunnen sporen en processen betrouwbaarder en efficiënter kunnen maken. Denk bijvoorbeeld aan het opsporen van creditcardfraude aan de hand van typische patronen in transactiedata. Er is echter ook een keerzijde aan deze ontwikkeling. Door de nadruk op compliance worden afwijkingen als negatief gezien. Er kunnen echter ook positieve afwijkingen zijn waarvan we veel kunnen leren!

De “Positive Deviance” (PD) beweging is gebaseerd op de observatie dat er in elke gemeenschap individuen of groepen zijn die op onconventionele wijze problemen oplossen en juist effectief zijn door zich niet te conformeren aan de “normale” gang van zaken. De beweging is ontstaan in de jaren zeventig toen bij onderzoek opviel dat sommige kinderen in een achterstandssituatie het veel beter deden dan andere kinderen in vergelijkbare condities. Bij PD gaat het om het mobiliseren van veelal onzichtbare en onconventionele oplossingen. Het idee is reeds toegepast in onderwijs, gezondheidszorg en voeding. Denk aan oplossingen om meer meisjes techniekonderwijs te laten volgen, het aantal MRSA besmettingen in ziekenhuizen terug te dringen, en ondervoeding te voorkomen.

De beschikbaarheid van “Big Data” maakt het mogelijk ook economisch voordeel te verkrijgen door slim te leren van positieve afwijkingen. Denk aan de bakkerijen in de 850 winkels van Albert Heijn. Waarom verkopen sommige bakkerijen meer brood dan anderen? Dit hangt natuurlijk af van de locatie en grootte van de winkel. Het doel is echter de minder zichtbare en soms onconventionele factoren te bepalen die ervoor zorgen dat sommige winkels veel meer brood verkopen dan verwacht.

Vaak zal het aantal positieve afwijkingen erg klein zijn. Daarom voldoen klassieke data mining technieken zoals het afleiden van beslisbomen (decision tree learning) niet. Deze laten meestal de typische te verwachten afhankelijkheden zien. Ook is het aantal mogelijke verklaringen vaak veel groter dan het aantal gevallen waar we iets

van kunnen leren. Daarom is het zaak juist de *onverwacht* positieve afwijkingen op te sporen en deze gevallen in detail te onderzoeken. Zie www.positivedeviance.org voor voorbeelden.

De “data scientist” van de toekomst moet oog hebben voor weinig voorkomende positieve uitzonderingen. Nederland kan deze tegendraadse en non-conformistische individuen en groepen prima gebruiken!

Biografie

Prof.dr.ir. Wil van der Aalst is als hoogleraar verbonden aan de Technische Universiteit Eindhoven (TU/e). Internationaal wordt hij gezien als de expert op het gebied van Business Proces Management (BPM). Hij heeft diverse doorbraken op zijn naam staan op gebieden uiteenlopend van procesmodellering en simulatie tot workflow management en process mining. Volgens Google Scholar is hij een van de meest geciteerde informatici in de wereld. Ook heeft hij aan de wieg gestaan van belangrijke open-source initiatieven zoals YAWL en ProM.