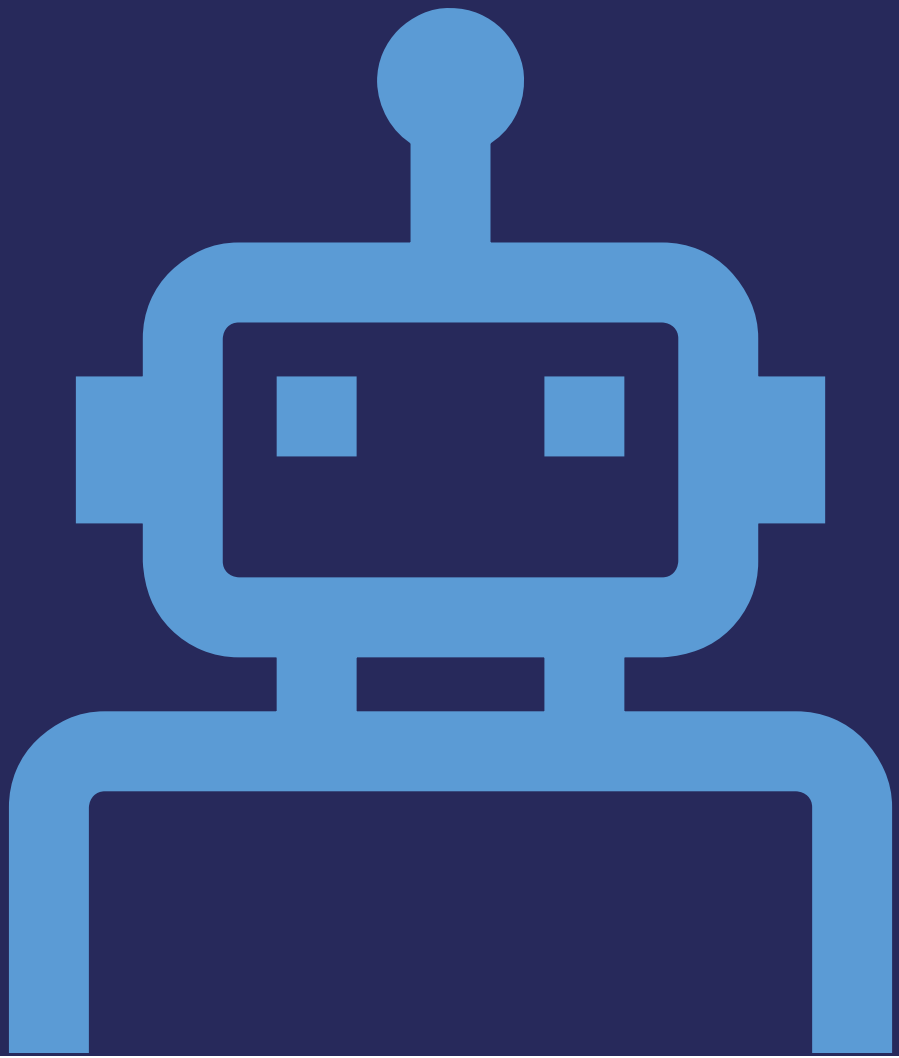
Trendskout

Automated Machine Learning voor Business Intelligence

Trendskout Tech Labs Whitepaper



Inhoudsopgave

3 Inleiding

4 Automated Machine Learning

5 Van start met Machine Learning

7 Datavoorbereiding en -transformatie

8 Een iteratief proces

10 IT als profit center - Conclusie

Inleiding

De groeiende wereldwijde vraag naar machine learning capaciteit is groter dan het aanbod. Om de kloof tussen vraag en aanbod te overbruggen, stelt Automated Machine Learning een nieuwe klasse datawetenschappers in staat om aan deze vraag naar capaciteit te voldoen: de **citizen data scientist**.

Gartner[1] definieert een citizen data scientist als een persoon die geavanceerde diagnostische, voorspellende of prescriptieve modellen ontwikkelt waarvan de eigen primaire rol buiten het veld ligt van statistieken of analyses. Veel data-analisten en business intelligence professionals zullen waarschijnlijk doorgroeien naar rollen als citizen data scientists. Automated Machine Learning is voor hen een ideale oplossing om die sprong te wagen.

Automated Machine Learning: extra tool in de toolkit

Met Automated Machine Learning hoeft u geen datawetenschapper van opleiding te zijn om onmiddellijk voorspellende inzichten te genereren. Als je tools gebruikt zoals Tableau, Power BI of Excel heb je vaak al meer dan genoeg technische bagage op zak om met Automated Machine Learning aan de slag te gaan.

De volgende fundamentele verschuiving in de evolutie van analytics is al aan de gang. De volgende generatie geautomatiseerde analysetechnologieën veranderen de manier hoe analyses worden benaderd. Net zoals self service BI oplossingen als Power BI een grote impact hadden op Business Intelligence, zal Automated Machine Learning het spel opnieuw veranderen.

> "Tegen 2023 zal 50% van de analytische vragen automatisch worden beantwoord of via Natural Language Processing."

Early adopters van Automated Machine Learning prijzen de ongekende snelheid om inzicht en een groter concurrentievoordeel te verwerven. Gartner voorspelt voor 2023 dat augmentatie 2,8 miljard euro aan bedrijfswaarde zal creëren en 6,2 miljard uur arbeidsproductiviteit.

Automatiseren van repetitieve taken van Machine Learning

Automatisering is in deze geen reden tot ongerustheid. Deze oplossingen zijn gericht om de kennis van data te combineren met de kracht van Machine Learning. Waar business intelligence en visuele analyse het toelaten om te begrijpen wat er in het verleden is gebeurd, laat Automated Machine Learning u navigeren door de toekomst. Of anders gezegd, Automated Machine Learning automatiseert de repitieve taken van Machine Learning, en laat IT teams toe om zich te richten op innovatie en het verstevigen van hun positie als profit center in de organisatie.

[1] https://www.gartner.com/newsroom/id/3570917

Automated Machine Learning

Trendskout automatiseert het machine learning proces van data ingestie tot operationalisering van het model. Het levert directe waarde en ongeëvenaard gebruiksgemak - geen ingewikkelde wiskunde of scripting vereist.

Sleutelinzichten en verborgen patronen vinden

Trendskout automatiseert o.m. feature-engineering, en het vinden van sleutelinzichten en verborgen patronen. De onderliggende technologie versnelt analytisch onderzoek van miljoenen variabele combinaties die te veel tijd zouden vergen voor handmatige menselijke analyse.

Zelfstandige hypertuning versnelt proces

Toepassingen voor zowel AI als Deep Learning worden gestuurd door algoritmes, data en parameters, die alle drie nauwkeurig samenwerken voor de beste prestaties. De keuze en wisselwerking tussen algoritme, dataverwerking en parametrisatie – ook wel hypertuning genoemd – wordt door de Trendskout AutoML zelfstandig uitgevoerd. Dat versnelt het volledige proces: van conceptualisering tot oplevering van een performante toepassing voor AI en Deep Learning.

Hiervoor werkt Trendskout met alle geavanceerde Machine- en Deep Learning frameworks zoals Tensorflow, XGBoost, Apache Spark, Theano, Scikit Learn en meer. Dit ondersteund door een **visuele drag & drop editor**

- Ingestie van databronnen,
- Selecteren van doelvariabelen om te voorspellen,
- Automatisch genereren van feature extractie, feature selectie rapporten, en het zoeken door tal van mogelijke datatransformatie-algoritme-hypertuning combinaties door ons Genetisch Algoritme,
- Snel operationaliseren van modellen.

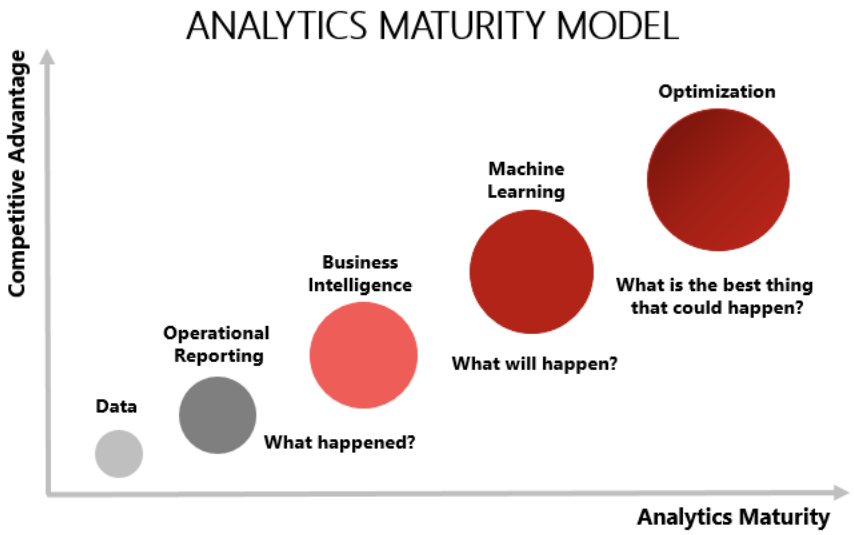
Tijd verkorten tussen inzicht en actie

Machine Learning ontwikkelingen die vroeger weken of maanden in beslag namen, kunnen nu worden afgerond in uren. Door het invoegen van Trendskout Automated Machine Learning in de bestaande reporting of processen, wordt de tijd tussen inzicht en actie drastisch verkort.

In deze whitepaper bespreken we o.m. hoe aan de slag te gaan met Automated Machine Learning. We introduceren verschillende Machine Learning concepten, geven belangrijke data preparatie tips, geven een stappenplan hoe Trendskout kan toegepast worden op uw data. Tot slot leggen we het verband met de rol van een modern IT team en deze evoluties.

Van start met Machine Learning

Laten we, voordat we ingaan op Machine Learning, even kijken naar het analytics maturity model en hoe het business intelligence en analytics aanvult.



Opvragen, samenvoegen en voorbereiden gegevens

De volwassenheid van analytics begint met ruwe data omzetten in rapportages om op intuïtie gebaseerde besluitvorming te verminderen. Gegevens worden geaggregeerd in datawarehouses. Op dit niveau van analytische volwassenheid wordt de meeste tijd besteed aan het opvragen, samenvoegen en voorbereiden van gegevens. Er blijft weinig tijd over voor data-analyse.

Analyse van bevindingen uit het verleden

In de volgende fase wordt business intelligence gebruikt om te begrijpen wat er is gebeurd in het verleden met analytische olap-cubes, draaitabellen, datavisualisatie en dashboards. Deze essentiële datagestuurde beslissingstools vereisen doorgaans handmatige ontwikkelingsprocessen om tellingen, sommen en gemiddelden te aggregeren. Bevindingen weerspiegelen vaak één enkel punt in de tijd en houden doorgaans geen rekening met statistische significantie.

Vooruitblikken met Automated Machine Learning

Machine learning wordt gebruikt om te voorspellen wat er gaat gebeuren. Machine learning biedt systemen de mogelijkheid om automatisch te leren van ervaringen uit het verleden zonder expliciet geprogrammeerd te zijn. Machine learning legt statistisch significante relaties tussen invoer, vaste variabelen en doelvariabelen. Automated Machine Learning versnelt het proces.

Machine Learning: verborgen patronen en trends detecteren

Machine Learning kan op efficiënte wijze grote hoeveelheden gegevens analyseren die te veel variabelen bevatten voor handmatige business intelligence of traditionele statistische analyse technieken. Tijdreeksen, regressie, neurale netwerken, beslissingsbomen, classificatie, random forests, support vector machines en vele andere soorten machine learning-algoritmen kunnen automatisch het signaal in de data oppikken. Verborgen patronen en trends in gegevens die de menselijke geest niet kan detecteren, worden geïdentificeerd.

Use-cases van Automated Machine Learning

Veelvoorkomende use-cases voor het toepassen van Automated Machine Learning zijn onder meer: financiële credit scoring, churn voorkomen, voorraad optimaliseren, direct marketing, conversiemodellering, prijsbepaling, prognoses, fraudedetectie en risico's beheersen.

> "Machine Learning vervangt Business Intelligence niet, maar vult het aan."

Voordelen van Automated Machine Learning

- Bespaart tijd door snel miljarden combinaties te evalueren.
- Vindt het signaal in de ruis, legt verborgen patronen en trends bloot die voor het menselijk oog onzichtbaar zijn.
- Verbetert de accuraatheid door het toepassen van Deep Learning technieken die voor een manuele aanpak niet haalbaar zijn.
- Vermindert de *bias* – vooroordeel – in besluitvorming door de pure data gedreven inzichten vs menselijke interpretatie van rapporten.

Business intelligence skills zijn een solide basis voor het toepassen van Automated Machine Learning, en versnellen het toepassen van waardevolle inzichten. Domeinkennis, probleemoplossend denken en data zijn de hoeksteken van succes.

Datavoorbereiding en -transformatie

Datacollectie en – voorbereiding is een cruciale stap in ieder Machine Learning project. Trendskout biedt tal van technieken en automatisaties om dit proces zo efficiënt mogelijk te laten lopen, maar het begrijpen van de betekenis van data voor uw organisatie is niet mogelijk voor AI. De juiste data selecteren is immers nauw verbonden met het kiezen van goede optimalisatiedoelen.

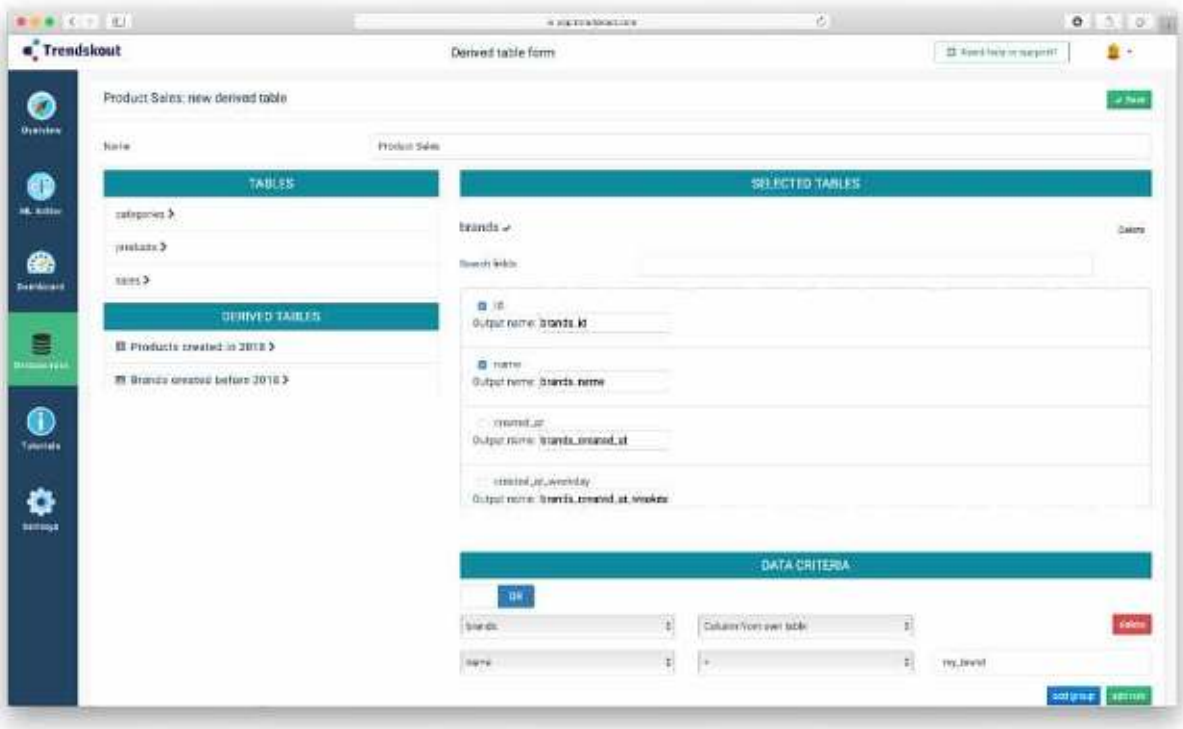
Datasources linken aan elkaar met behulp van dataprepper

Als de data geselecteerd is, kan deze snel toegevoegd worden als datasource in ons systeem. Dit kunnen relationele databanken, tekstbestanden, NoSQL databanken, datalakes of combinaties daarvan zijn.

> "Trendskout automatiseert, naast het vinden van het beste model, ook de nodige datatransformatie van uw data."

Via de ingebouwde dataprepper kunt u datasources linken aan elkaar, en zo semantische betekenis geven aan uw data die dan gebruikt wordt door de Automated Machine Learning motor.

Dit betekent niet dat u zelf aan de slag moet om allerlei scripting uit voeren, uw data wordt ontsloten met een Data Prepping wizard die zelf alle nodige technische transformatie en opschoning uitvoert.



Een iteratief proces

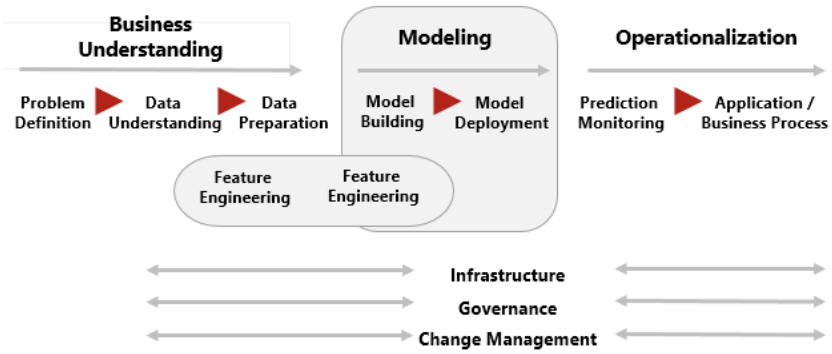
Vergelijkbaar met Kimball-dimensionale modellering die wordt gebruikt in business intelligence-projecten, is het proces voor het ontwikkelen van hoogwaardige Machine Learning-modellen iteratief van aard.

Stakeholders, workflow en integratie met bedrijfsprocessen

Start nooit een project zonder eerst na te denken over stakeholders, de workflow en hoe uw machine learning-oplossing wordt geïntegreerd in bedrijfsprocessen om een positieve impact te maken. Een best practice is om te beginnen met kleine, quick win projecten om vaart te krijgen en verder te bouwen op het momentum.

Identificeren van kansen en goede uitvoering essentieel voor succes

Succes ligt in het identificeren van kansen en het uitvoeren van zoveel mogelijk ervan. U kunt aan de slag door gegevens te gebruiken die al beschikbaar zijn, immers gegevensbronnen zullen nooit volledig of perfect zijn. Moderne gegevensvoorbereiding en Automated Machine Learning stellen u in staat om snel te schakelen.



Probleemdefinitie, doelstellingen en eisen bepalen

Het Machine Learning-proces begint met domeinkennis. Deze eerste belangrijke stap richt zich op het definiëren van het juiste probleem om op te lossen, het herkennen van de business doelstellingen en eisen. Na het selecteren van een probleem, start het verzamelen van gegevens. Tijdens deze stap raak je vertrouwd met de beschikbare databronnen, identificeer je problemen met de gegevenskwaliteit en voer je verkennende analyses uit. Tijdens de data preparation worden gegevens verder verwerkt.

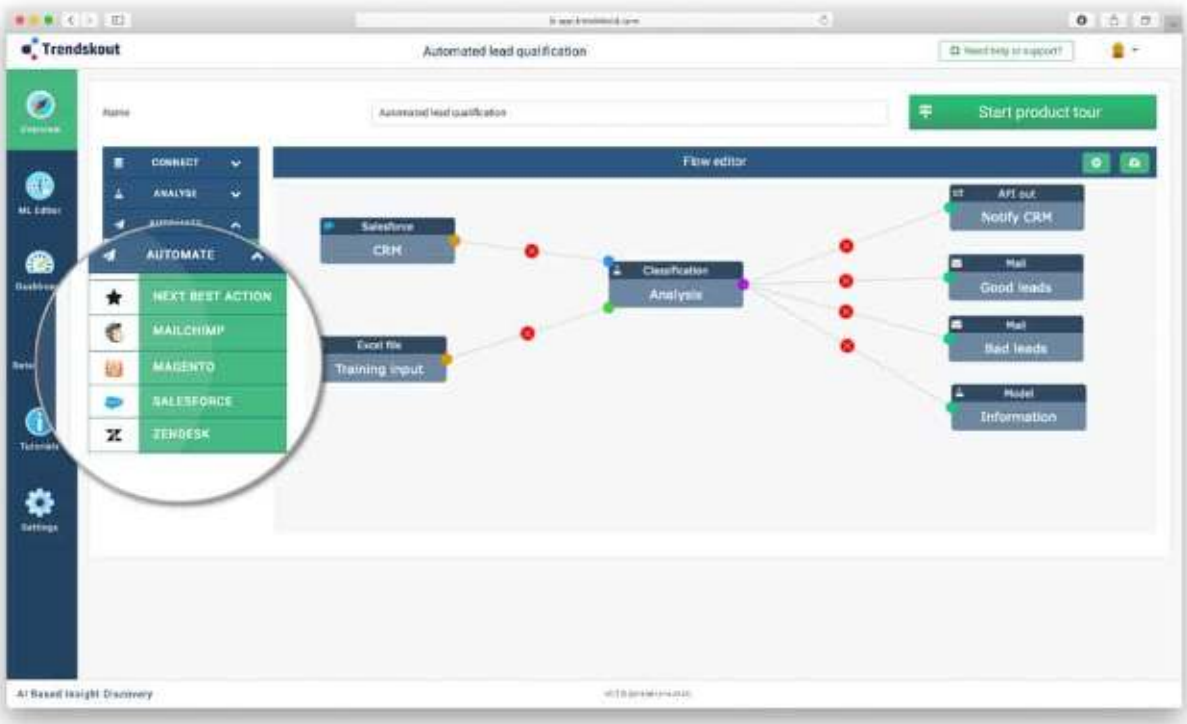
> "Data verandert voortdurend. Drift detectie en automatische model updates zitten standaard in het Trendskout platform."

Tijdens het modelleren selecteert u een doel om te voorspellen en start het Automated Machine Learning platform automatisch met het trainen van honderden voorspellende modellen, ze te scoren en te rangschikken op geselecteerde kwaliteitsmetrieken.

Resulterend model operationaliseren

Nadat Trendskout de juiste combinatie van datatransformaties, algoritme selectie en hypertuning heeft gevonden bent u klaar om het resulterende model te operationaliseren. Hiervoor kan u kiezen uit verschillen Automate-acties die het toepassen van uw model vereenvoudigen:

- Verschillende plug-ins voor courante systemen
- API endpoints
- Database connecties
- FTP
- NoSQL & Datalakes export
- …



Conclusie

IT-professionals nemen al decennia het voortouw in innovatie. Machine Learning en aanverwante technologieën brengen tal van opportuniteiten om deze rol verder te verstevigen. Om deze kansen te verzilveren is het voor IT-teams cruciaal te focussen op de juiste toegevoegde waarde.

Co-creatie proces

Het ontsluiten van deze waarde zit hem in een co-creatie proces met de business user om een predictief vraagstuk te beantwoorden, met een IT-team dat als citizen datascientist en dataprofessional de business user gidst in het datalandschap.

Vertragende taken automatiseren

Bij Trendskout is het onze ervaring dat het uitvoeren van tijdrovende en complexe taken om een Machine Learning model te creëren weinig toegevoegde waarde brengt aan dit co-creatie proces en vaak als vertragend ervaren wordt. Door het gebruiken van Automated Machine Learning kunnen deze taken geautomatiseerd worden en kan het IT-team zichzelf verder profileren als profit center.

Wilt u ook ontdekken wat de mogelijkheden zijn voor uw organisatie? Neem dan direct contact met ons op voor een gratis demo!

Ontdek hoe onze AI Sofware werkt

EN MEER

Trendskout ondersteunt een brede waaier aan mogelijkheden in verschillende sectoren naast die hierboven vermeld. Aarzel dus niet om contact met ons op te nemen om te ontdekken hoe Trendskout directe toegevoegde waarde kan creëren voor uw organisatie.

+3293934628

hello@trendskout.com

https://trendskout.com/

Trendskout