



Data gestuurde graafschade inspectie



Agentschap Telecom
Ministerie van Economische Zaken

Samenvatting

Als onderdeel van het Ministerie van Economische Zaken en Klimaat is het Agentschap Telecom behalve als uitvoerder en toezichthouder van wet- en regelgeving op gebied van telecommunicatie, ook verantwoordelijk voor het toezicht op de Wet informatie-uitwisseling bovengrondse en ondergrondse netten en netwerken (WIBON). Doel van deze wet is het voorkomen van gevaar of economische schade door beschadiging van ondergrondse kabels en leidingen. Jaarlijks voert het Agentschap Telecom honderden inspecties uit.

Om de hoeveelheid glasvezel gerelateerde graafschades te reduceren startte het Agentschap Telecom in 2020 een pilot om informatie uit een aantal relevante databronnen (cartografische data, gebiedspolygoenen, twitter informatie enz.) op interactieve wijze via een Shiny app beschikbaar te stellen aan de inspecteurs, ter ondersteuning bij het uitvoeren van hun inspecties.

Daarnaast is er een Machine Learning model ontwikkeld dat op basis van historische data de kans op graafschade kan voorspellen op het moment dat er een KLIC-melding (graafmelding) bij het Kadaster plaatsvindt. Deze voorspelling is ook beschikbaar in de Shiny app.

Vanwege de goede inhoudelijke resultaten van het voorspelmodel hebben Agentschap Telecom, Full Orbit, DICTU en Oracle in een vervolgstap de handen ineengeslagen om te kijken of het mogelijk was om door gebruik van OML4R, in-database Machine Learning technologie van Oracle, de huidige opzet van het van het model efficiënter te maken met het oog op toekomstige operationalisering. De resultaten hiervan waren veelbelovend, zowel met betrekking tot reductie van de doorlooptijd van het trainen van het Machine Learning algoritme, als ook het elimineren van handmatige stappen bij het beschikbaar stellen van nieuwe data.

In de toekomst zal worden bekeken hoe het voorspelmodel verder kan worden geoperationaliseerd en geïmplementeerd in de werkprocessen, zodat het daadwerkelijk bij de uitvoering van de graafinspecties ingezet kan worden.



Full Orbit
DRIVE VALUE WITH DATA





Over Agentschap Telecom

Agentschap Telecom is een onderdeel van het Ministerie van Economische Zaken en Klimaat. Het is zowel uitvoerder als toezichthouder van wet- en regelgeving op gebied van telecommunicatie inclusief Caribisch Nederland (Bonaire, Saba en Sint Eustatius).

Het agentschap doet dat ook op andere terreinen, zoals bijvoorbeeld de wet die toeziet op veilig graven (Wibon), weegschalen, goud, zilver of platina en ruimtevaartactiviteiten.

WIBON

Agentschap Telecom houdt toezicht op de Wet informatie-uitwisseling bovengrondse en ondergrondse netten en netwerken (WIBON). Doel van de WIBON is het voorkomen van gevaar of economische schade door beschadiging van ondergrondse kabels en leidingen.

Het voorkomen van graafschades is een taak van opdrachtgevers, grondroeders (degene onder wiens verantwoordelijkheid of leiding graafwerkzaamheden worden verricht) en netbeheerders, en gebeurt door het delen van informatie en het maken van afspraken. In het kader van deze wet worden er door het Agentschap Telecom jaarlijks honderden inspecties uitgevoerd op de ruim 700.000 mechanische graafwerkzaamheden die er plaatsvinden.

Graafschade

Ondanks de door deze wet beoogde informatie-uitwisseling, veroorzaakt graafschade veel (maatschappelijke) overlast en heeft het hoge directe herstelkosten, ruim 38€ miljoen in 2020. Bovendien kan het in sommige gevallen leiden tot gevaarlijke situaties. Dit is de laatste jaren met name het geval bij de aanleg van glasvezel waarbij gasleidingen worden geraakt.



¹ Selectie Infographic "WIBON en schade door graven" Agentschap Telecom 31-5-2021

Iedereen die machinale graafwerkzaamheden uitvoert is verplicht om, vóór aanvang van de werkzaamheden, een zogenaamde KLIC-melding (graafmelding) bij het Kadaster te doen, inclusief de opgave van het graafpolygoon, een plattegrond met de plek waarop gegraven gaat worden. Het Kadaster haalt vervolgens gegevens over de ligging van de ondergrondse netwerken in het graafpolygoon bij de betrokken netbeheerder(s) op en stuurt deze gebiedsinformatie over het opgegeven graafgebied naar de aanvrager van de melding. De gebiedsinformatie bevat o.a. de datum waarop de graafwerkzaamheden zullen starten, de gegevens van de grondroerder en het graafpolygoon waarbinnen zal worden gegraven.

“ De hoge veroorzaakte economische schade in combinatie met het kunnen beschikken over een groot aantal relevante databronnen, maakte glasvezel gerelateerde graafschade voor ons een relevant thema voor het starten van een pilot. De doelstelling hiervan was om met behulp van data-analyse het aantal meldingen van graafschade te kunnen reduceren ”

Dr. Rianne van Dalen

Data Scientist
Agentschap Telecom

Stap 1: Pilot reductie graafschades

Om de hoeveelheid graafschades te reduceren is het Agentschap Telecom in 2020 een pilot gestart om een aantal relevante informatiebronnen met betrekking tot de aanleg van glasvezel samen te voegen en deze weer te geven in een interactieve webapplicatie op basis van het R Shiny framework. Via deze app is alle relevante data te allen tijde door de inspecteurs te benaderen via de laptop, tablet of mobiele telefoon ten behoeve van het uitvoeren van inspecties.

Data-elementen

De volgende data-elementen worden in de Shiny app samengevoegd:

- Actuele kaartgegevens via PDOK (Publieke Dienstverlening op de Kaart). Op PDOK stelt de overheid open datasets beschikbaar met actuele geo-informatie.



- Informatie m.b.t. de actuele aanleg van glasvezel in heel Nederland middels via het Kadaster verkregen gebiedspolygonen.
- Luchtfoto's en bodemkaarten voor visualisatie
- De GPS-locatie van de inspecteur. Eén van de wensen van de inspecteurs was om in het veld alle actuele graafmeldingen in de buurt in één oogopslag te zien. Het kan voorkomen dat na het uitvoeren van een inspectie, een andere KLIC-melding in de buurt actief is. Zo kan efficiënter toezicht worden gehouden op graafactiviteiten.
- Actuele twitterberichten over glasvezel. De onafhankelijke site Glasvezelgids.nl informeert over de ontwikkelingen op het gebied van glasvezelinternet voor consumenten en bedrijven in Nederland. Deze informatie is voor de inspecteurs relevant, omdat ook berichten vanuit de maatschappij aanleiding kunnen zijn voor nadere inspectie.

Voorspelmodel

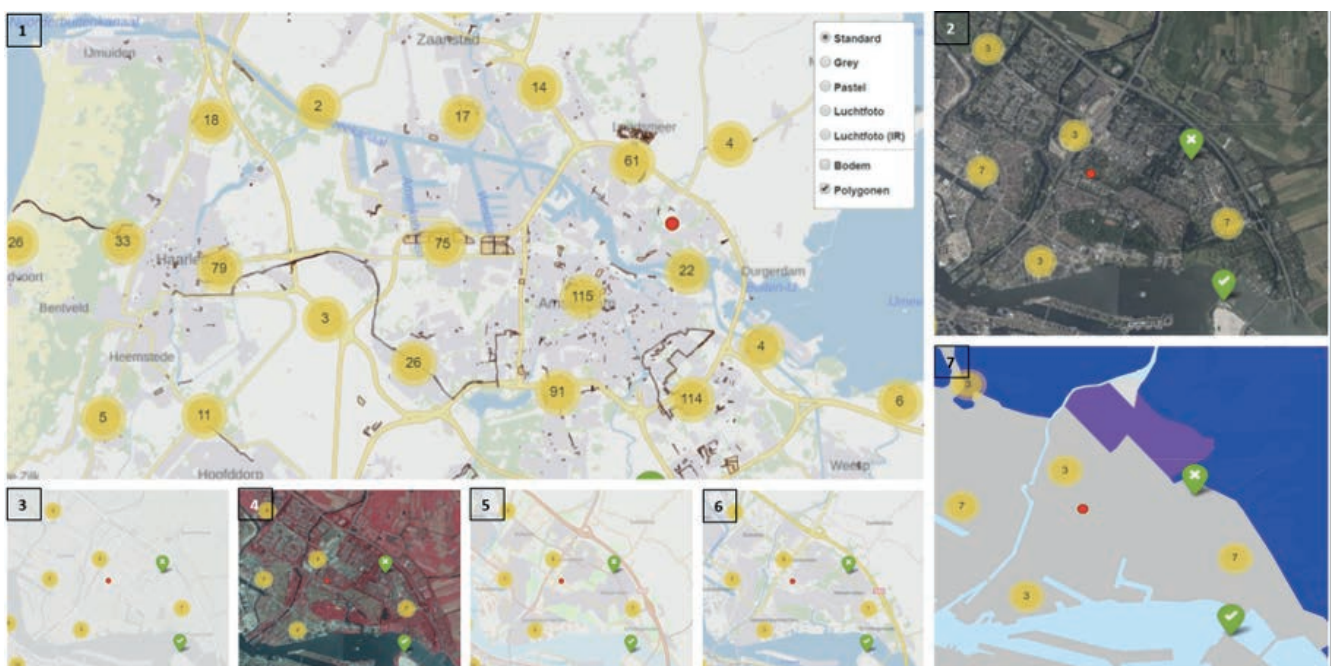
Naast de ontwikkeling van bovengenoemde Shiny app wilde Agentschap Telecom ook onderzoeken of er een Machine Learning model kon worden ontwikkeld dat, op basis van de informatie die is opgegeven tijdens het aanvragen van een KLIC-melding, de kans op graafschade kan voorspellen.

Bij de ontwikkeling van dit Machine Learning model zijn er vijf classificatie voorspeltechnieken toegepast op historische kadasterdata met verschillende parameters. Uiteindelijk is het "random forest" algoritme gebruikt waarbij de interpreteerbaarheid van het voorspelmodel (uitleg aan derden) in combinatie met de voorspellende waarde de doorslag heeft gegeven.

De belangrijkste voorspellers in het model zijn:

- Percentage graafschade per grondroerder (aantal graafmeldingen met schade/ totaal aantal graafmeldingen);
- Aantal coördinaten van het graafpolygoon;
- Aantal thema's van de graafmelding (warmte, water, datatransport, etc.);
- Aantal en type werk van de graafmelding (heien, persing/boring, bodemsanering, etc.).

Uiteindelijk bleek de accuraatheid van het model hoog, 82% van alle voorspellingen is correct. Ook is de PPV (proportie aantal voorspelde graafmeldingen met schade die correct zijn) vrij hoog: 62% van het aantal correct voorspelde schades is ook daadwerkelijk een schade. Dit is een belangrijke indicator voor 'risico-gestuurd' toezicht. Als inspecteurs in de toekomst gebruik gaan maken van de uitkomsten van voorspelmodellen om hun werk te prioriteren, dan wil je graag dat ze naar graafmeldingen gaan



2 Voorbeeld van een aantal data elementen die in de Shiny app worden weergegeven



waarbij de kans klein is dat ze voor “niets” komen.

De uitkomst van het voorspelmodel is een kans op graafschade voor elke graafmelding die tussen de 0 en 100% ligt.

Stap 2: Vergroten efficiency voorspelmodel

Agentschap Telecom heeft de pilot uitgevoerd op een gedeelde server met R en RStudio om de data te modelleren en het random forest algoritme aan te sturen. Vanwege de goede inhoudelijke resultaten werd ervoor gekozen om een vervolgstap te zetten in de vorm van een POC, met als doel te kijken of het voorspelmodel efficiënter kan worden gemaakt met het oog het operationaliseren in te toekomst. Aandachtspunten bij dit POC waren:

- het reduceren van de doorlooptijd die nodig is om het random forest algoritme te trainen. In de oorspronkelijke opzet van de pilot, op basis van opensource technologie, duurde het trainen van het model al snel 6 uur. Met de in de toekomst verwachte toename in data zal dit alleen maar nog langer gaan duren.
- kijken of er een opzet mogelijk is waarbij nieuwe data gemakkelijk beschikbaar kan worden gesteld zonder handmatige stappen als export en lokale opslag.

Om het voorspelmodel van Agentschap Telecom verder te optimaliseren en in de toekomst te operationaliseren hebben Agentschap Telecom, Full Orbit, Oracle en de Dienst ICT Uitvoering (DICTU) de handen ineengeslagen om via een POC te experimenteren met het uitvoeren van in-database Machine Learning op basis van Oracle Machine Learning for R (OML4R).

Wat is OML4R?

Oracle Machine Learning for R kan gezien worden als extensie op 'opensource R' waarbij R aangevuld wordt met Oracle R. Naast alle open source mogelijkheden die R al heeft, biedt OML4R diverse packages voor data manipulatie, data visualisatie, statistiek en in-database Machine Learning met data vanuit een Oracle database. Het grote voordeel is dat je de rekenkracht van de database kunt gebruiken, én de

database is natuurlijk een ideale plek om te voorzien van nieuwe data die dan direct beschikbaar zijn voor een voorspellend model.

Graafschade voorspellen met OML4R

Om dit alles te realiseren hebben bovengenoemde partijen allereerst gekeken of het mogelijk was om de voorspellingen van graafschade te doen middels OML4R. Full Orbit, DICTU (verantwoordelijk voor de database servers van Agentschap Telecom) en Oracle hebben gezamenlijk de installaties begeleid/uitgevoerd en gekeken naar de impact van installatie en gebruik van OML4R binnen het bestaande Oracle Exadata database-landschap.

De resultaten waren veelbelovend! De mogelijkheden tot in-database data manipulatie en preparatie middels de Oracle R packages waren een toereikend alternatief voor de opzet waarbij gebruik gemaakt werd van opensource R package tidyverse. De vervolgstap, het uitvoeren van het in-database random forest algoritme gaf inhoudelijk vergelijkbare resultaten ten opzichte van de initiële opzet. Het grote voordeel zat daarnaast in de doorlooptijd van het trainen van het random forest model, in de huidige opzet duurt dit 6 uur terwijl de doorlooptijd bij een vergelijkbare dataset teruggebracht werd tot 2 uur middels gebruik van OML4R.

“Onze doelstelling bij het uitvoeren van de POC was om te zien of we de performance van het voorspelmodel konden verbeteren, het aantal handmatige stappen konden reduceren en te bezien of het model in ons BI landschap kan worden geïncorporeerd”

Mustafa Bozkurt

Senior Data Consultant
Agentschap Telecom

De toekomst

Op basis van de goede resultaten van het voorspelmodel zal de vervolgstap bestaan uit het bekijken op welke wijze het model daadwerkelijk geoperationaliseerd kan worden. Aandachtspunten hierbij zijn de technologische integratie van het model in het bestaande Business Intelligence landschap, alsmede het





Agentschap Telecom
Ministerie van Economische Zaken

creëren van voldoende draagvlak bij de beoogde eindgebruikers, zodat zij de uitkomsten van het model durven te vertrouwen en bij het uitvoeren van hun inspecties zullen gebruiken.

Meer weten

Wil je weten hoe ook jij meer impact kunt maken met Data Science binnen je bedrijf? Neem dan gerust contact met ons op. We vertellen er graag meer over.

Contact opnemen

Full Orbit

Tingietersweg 90
2031 ES Haarlem
E info@full-orbit.com
T +31 (0)20 5300 390

Socials:



Full Orbit
DRIVE VALUE WITH DATA