

Observatie Inzichten Handleiding

Inleiding

Observatie Inzichten is een functie in de Sfeer App die je helpt betere, betrouwbaardere kaarten te maken. In de praktijk merken we dat de kwaliteit van een model niet alleen afhangt van hoeveel observaties je toevoegt, maar vooral van *welke* observaties je toevoegt. Observatie Inzichten maakt dit inzichtelijk met een set duidelijk interpreteerbare indicatoren van observatiekwaliteit, uitgedrukt in waardes waar je meteen op kan handelen. Hiermee zie je sneller waar je model sterk is en waar je met kleine aanpassingen grote winst kunt behalen.

Het bouwen van een goed model is vaak een iteratief proces: observaties toevoegen, resultaat bekijken, bijsturen en opnieuw trainen. Observatie Inzichten geeft structuur aan dat proces door signalen te geven over de datakwaliteit en dekking van je observaties. Zo kun je gericht werken en vergroot je de kans dat het model een stabiele en betrouwbare kaart maakt.

Wat is de functie van Observatie Inzichten?

Observatie Inzichten vertaalt je observaties in indicatoren. Het helpt je te begrijpen waarom een model goed of minder goed presteert, en welke aanpassing waarschijnlijk het meeste effect heeft.

Hoe kan je het gebruiken?

Gebruik Observatie Inzichten als een checklist tijdens het verbeteren van je model:

1. Begin met een bestaand model.
2. Bekijk de indicatoren en let op waarschuwingen of afwijkende waarden.
3. Pas gericht aan: elke indicator wijst je naar een specifiek type probleem in je observaties.
4. Train opnieuw en vergelijk: zie of de indicatoren én de kaartkwaliteit verbeteren.
5. Herhaal indien nodig tot de kaart stabiel en goed is voor jouw doel.

Waar moet je voor oppassen bij het gebruiken?

Observatie Inzichten is een hulpmiddel, geen automatische “waarheid”. Blijf altijd zelf nadenken en houd rekening met het volgende:

- **Bekijk indicatoren altijd samen.** Eén indicator kan er “goed” uitzien terwijl een andere een groot probleem aanwijst. Bijvoorbeeld: een groot oppervlak aan observaties is minder effectief als alle observaties uit één jaar komen of één klasse domineert.
- **Observatiekwaliteit is meer dan positieve indicatoren.** Naast de indicatoren die wij geven, tellen ook inhoudelijke aspecten mee, zoals:
 - **Zuiverheid/consistentie van labels:** zijn observaties echt representatief voor die klasse?

- **Variatie binnen een klasse:** dekt je klasse verschillende situaties (bodem, vocht, beheer, seizoen)?
- **Representativiteit:** lijken je observaties op de gebieden waar je straks voorspellingen voor wilt maken?
- **Groen is geen garantie.** Als alle signalen op groen staan betekent dat nog niet dat de kaart goed is. Observaties kunnen technisch in orde zijn maar inhoudelijk onzuiver of tegenstrijdig. Ook kan het zijn dat bepaalde natuur simpelweg moeilijk te onderscheiden is voor het model.

Gebruik Observatie Inzichten dus als kompas: het wijst je de richting, maar jij bepaalt de koers op basis van domeinkennis en het doel van je kaart.

Kom je er niet uit? Onze support team helpt je graag verder (support@spheer.ai).

Inhoudsopgave

Inhoud

Inleiding	1
<i>Inhoudsopgave</i>	2
Totaal observatieoppervlak	4
Klassenbalans (voor classificatie modellen)	6
Waardebalans (voor regressie modellen)	7
Temporele dekking	8
Observatiegrootten	10

Totaal observatieoppervlak

Wat is het?

Het totaal observatieoppervlak is de som van het oppervlak van alle ingetekende observaties. Hierbij worden observaties die voor meerdere jaren gelden voor ieder jaar meegeteld.

Voorbeeld:

In dit voorbeeld is het totaaloppervlak van de observaties 8.4 ha, waarmee het in de lichtoranje zone ligt. De groene lijn ligt vanaf de 10 tot 500 ha, wat het oppervlak is waarin dit gebruikersmodel het beste werkt. In de lichtoranje zone kan actie ondernemen het model helpen, in de donkeroranje zone wordt actie ondernemen aangeraden.



Wat is het effect ervan?

Spheer gebruikt het oppervlak onder de observaties om van te leren. Het maakt voor het model geen verschil of dat oppervlak uit veel kleine observaties bestaat of één grote.

- **Te weinig oppervlak** geeft het model weinig informatie om van te leren, waardoor het moeite kan krijgen met voorspellen.
- **Te veel oppervlak**, kan ervoor zorgen dat nieuwe observaties weinig effect hebben op de voorspelling, aangezien er dan weinig ruimte overblijft om het model bij te sturen.

Daarnaast geldt een technische limiet: overschrijd je die, dan selecteert Spheer automatisch een subset van je observatiepixels.

Hoe kan je dit aanpassen?

Verhoog een lage score door meer observaties toe te voegen.

- **Tip:** gebruik hiervoor de volledigheidkaart, die toont waar het model nog extra observaties nodig heeft om een goede voorspelling te maken.

Verlaag een hoge score door (overbodige of dubbele) observaties te verwijderen of te verkleinen.

Als er goede redenen zijn om geen observaties te verkleinen of verwijderen (bijvoorbeeld bij een usecase met veel verschillende klassen of veel soorten van dezelfde klasse die individueel moeten worden ingetekend), kan je kiezen om een ander “Gebruikersmodel” te gebruiken. Er zijn gebruikersmodellen beschikbaar die beter omgaan met meer observatiedata. Voor meer

informatie over gebruikersmodellen raadpleeg de handleiding of vraag ons support team om hulp.

Hoe moet je met deze indicator omgaan?

Gebruik deze indicator vooral als grove richtlijn. Onze modellen zijn goed in het maken van kaarten vanaf een klein aantal observaties.

Klassenbalans (voor classificatie modellen)

Wat is het?

De klassenbalans doelt op de verhouding van (het oppervlak aan) observaties tussen de klassen in het model.

Voorbeeld:

In dit voorbeeld zien we 8 klassen. De balans laat zien dat de klasse “bos” met 39% veel beter vertegenwoordigd is dan bijvoorbeeld de klasse “zand(pad)”. Dit kan komen doordat dit gebied veel bos heeft of juist weinig zand. Het model kan hierdoor beter de klasse bos leren herkennen en meer moeite hebben met zand.



Wat is het effect ervan?

Spheer maakt kaarten gebaseerd op de informatie die het gegeven wordt. Als een klasse veel minder observaties heeft dan een andere, leert het model de tweede klasse een stuk beter te herkennen. Dit kan leiden tot een voorspelkaart van slechtere kwaliteit.

Hoe kan je dit aanpassen?

Vul de klasse met minder observaties aan door observaties toe te voegen, of verwijder of verklein observaties van de klasse met veel observaties.

Hoe moet je met deze indicator omgaan?

Het is niet nodig om de klassen precies te balanceren, er zijn vele goede redenen om meer of minder observaties te hebben tussen klassen onderling. Bijvoorbeeld als er veel meer van de ene klasse aanwezig is dan van de andere, of als een “overige” klasse gebruikt wordt waar veel verschillende voorbeelden voor ingetekend dienen te worden. Wel kan het nuttig zijn om bewust te zijn van grote verschillen. Als er zonder goede reden een onbalans is, is het de moeite waard om dit bij te stellen.

Waardebalans (voor regressie modellen)

Wat is het?

De waardebalans laat het waardebereik van je regressieobservaties zien. Door de balans kan je in één keer zien tussen welke regressiewaarden al je observaties zitten.

Voorbeeld:

In dit voorbeeld zien we dat de regressiewaarden tussen de 20% en de 70% liggen. Het model zal vooral waardes voorspellen die tussen deze twee waarden liggen.



Wat is het effect ervan?

Een regressiemodel voorspelt voornamelijk de waarden in het waardebereik, aangezien dit de voorbeelden zijn die het model gezien heeft. Door de balans kan je in één keer zien wat je bereik is en dus welke regressiewaarden moeilijker te voorspellen zullen zijn.

Hoe kan je dit aanpassen?

Als je waardebereik bepaalde waarden aan de uiteindes, zoals 0% en 100%, niet heeft en je wilt deze wel voorspellen, kan je observaties met deze waarden toevoegen.

Als je alleen regressiewaarden tot een bepaald niveau hebt, kan je er ook voor kiezen om je bereik te normaliseren. Je verandert dan je hoogste regressiewaarde naar 100% en doet dezelfde vermenigvuldiging met je andere waarden.

Een voorbeeld hiervan is:

1. Je hoogste regressiewaarde is 50%. Deze zet je naar 100% met de vermenigvuldiging:
 $100\% / 50\% = 2$.
2. Op alle andere waarden, bijvoorbeeld 0% en 30% doe je dezelfde vermenigvuldiging:
 - a. $0\% * 2 = 0\%$
 - b. $30\% * 2 = 60\%$
3. Op deze manier kan het hele regressiebereik benut worden

Hoe moet je met deze indicator omgaan?

Wanneer je alleen geïnteresseerd bent om waarden te voorspellen tussen de regressiewaarden die je hebt, of je hebt al een bereik tussen 0% en 100%, kan je deze indicator negeren.

Wanneer je wél geïnteresseerd bent in waarden buiten je bereik, voeg dan observaties met die waarden toe om je voorspellingen beter te maken.

Temporele dekking

Wat is het?

Temporele dekking is de verdeling van observaties tussen de jaren. Jaren verschillen, en dat heeft effect op de natuur. Bepaalde natuur kan er anders uitzien in een droog jaar dan in een nat jaar. Voor een goede vergelijking tussen jaren is het belangrijk dat het model voorbeelden ziet van de verschillende verschijningen van die natuur.

Voorbeeld:

In dit voorbeeld zien we een model met observaties in alle jaren. De temporele dekking indicator laat een “Goed” zien, omdat er in elk jaar een observatie zit.



Wat is het effect ervan?

De temporele dekking geeft een indicatie over hoe goed de vergelijking tussen jaren is. Wanneer er alleen in één jaar observaties toegevoegd zijn, kan het voor het model moeilijk worden om in een ander jaar een goede voorspelling te maken. Dit gebeurt wanneer dit jaar een afwijkend patroon laat zien, bijvoorbeeld het was veel droger dan andere jaren, wat een slecht beeld geeft voor andere jaren. Door in meer jaren observaties toe te voegen, wordt de stabiliteit over de jaren beter.

We weten uit eigen onderzoek dat van 1 jaar naar 2 jaren de grootste sprong in accuratesse geeft, gevolgd door van 2 jaren naar 3 jaren en daarna steeds minder. Als je een vergelijking tussen jaren wilt hebben, raden we altijd aan om minstens 2 jaren te gebruiken.

Hoe kan je dit aanpassen?

Vul het jaar met geen of minder observaties aan door observaties toe te voegen, of verwijder observaties van het jaar met veel observaties.

- **Tip:** In Spheer is het eenvoudig om voor een observatie aan te geven dat deze geldt voor meerdere jaren. Dit doe je tijdens het aanpassen van een observatie door bij de jaarselectie te klikken op “selecteer alle jaren”.

Zo is het niet nodig om voor ieder jaar unieke observaties te maken, het zal juist goed werken om een observaties te herbruiken voor verschillende jaren om zo variaties door de jaren te leren aan het model. Let er hierbij wel goed op dat de observatie ook accuraat is in ieder jaar, en snijd eventueel bij of verander zo nodig de klasse van een observatie.

Hoe moet je met deze indicator omgaan?

Als je een vergelijking tussen meerdere jaren wilt maken, zorg er dan voor dat er in meerdere jaren observaties zijn. Als je alleen geïnteresseerd bent in het jaar waar je de observaties

toegevoegd hebt, kan je minder aandacht aan deze indicator schenken, al zal ook in dat geval enige temporele dekking de kwaliteit van het model kunnen verbeteren.

Observatiegrootten

Wat is het?

Deze indicator toont het oppervlak per individuele observatie. Zo zie je in één oogopslag hoeveel observaties een goede grootte, groot of erg groot zijn.

Voorbeeld:

In dit voorbeeld zien we een model met in totaal 99 observaties, waarvan er veel klein zijn om genoeg variatie in de klassen weer te geven. Er zijn een paar observaties die aan de grote kant zijn, te zien in het lichtoranje gedeelte.



Wat is het effect ervan?

Erg grote observaties dekken vaak zoveel gebied dat veel informatie overbodig is. Meer informatie is niet slecht, maar veel observaties kunnen de traintijd van het model onnodig lang maken. Ook vergroten grote observaties de kans op ruis of onzuiverheid in deze observaties.

Zoals al eerder genoemd hebben wij een technisch limiet: overschrijd je die, dan selecteert Sfeer automatisch een subset van je observatiepixels.

Hoe kan je dit aanpassen?

Stem de grootte van observaties af op de variatie die ze bevatten. Grote observaties die steeds dezelfde structuur of hetzelfde type gebied beschrijven voegen weinig extra informatie toe: splits ze op in meerdere of maak ze kleiner. Kleine observaties zijn prima, maar zorg ervoor dat ze samen uiteenlopende omstandigheden dekken.

Hoe moet je met deze indicator omgaan?

Gebruik deze indicator in combinatie met de indicator voor het totaal observatieoppervlak voor het beste resultaat.