



Kwantificeren onzekerheid met 'Conformal Prediction'

Er zijn vele artikelen geschreven over het gebruiken van machine learning modellen voor actuariële doeleinden. Echter, de meeste van deze modellen produceren alleen een puntschatting, terwijl veelal een betrouwbaarheidsinterval of distributie gewenst is. 'Conformal prediction' is een techniek die de laatste jaren in opkomst en in ontwikkeling is, waarbij de onzekerheid gekwantificeerd kan worden voor ieder model type. In dit artikel beschrijf ik de beginselen van deze techniek.

ACHTERGROND

Er is een veelheid van machinelearningtechnieken beschikbaar, en de ontwikkeling van nieuwe technieken gaat gestaag verder. Bekende machinelearningtechnieken zijn lineaire / logistische / ridge / lasso regressie, decision trees, random forest, (extreme) gradient boosting en verschillende soorten neurale netwerken (zie Plat (2017)). Met uitzondering van de regressiemodellen leveren deze technieken geen kwantificering van onzekerheid op, althans niet met de juiste theoretische garanties.

De laatste jaren is conformal prediction in opkomst. Met deze techniek kan een betrouwbaarheidsinterval of distributie bepaald worden voor ieder mogelijk predictiemodel. Met de aanvulling van conformal prediction zijn machine learning modellen derhalve meer bruikbaar voor actuariële doeleinden.

INTRODUCTIE CONFORMAL PREDICTION

Conformal prediction is in verschillende vormen beschikbaar. De meest gebruikte heet inductive conformal prediction, waarvan de basisversie hieronder beschreven is.

Veronderstel een dataset $D = \{(x_i, y_i)\}_{i=1}^N$ met y_i de te verklaren variabele en x_i de verklarende variabelen. De eerste stap is om deze data te splitsen in drie verschillende sets: een (i) training, (ii) kalibratie en (iii) test set zodanig dat $D = D_{train} \cup D_{kal} \cup D_{val}$ met $N = n_{train} + n_{kal} + n_{val}$. De stappen waaruit conformal prediction bestaat zijn als volgt (zie bijvoorbeeld Anastasios et al (2022)):

1. Train een machine learning model op de training set D_{train} .
2. Definieer een maatstaf van onzekerheid $s(x, y)$, genaamd de 'non-conformity score function'. In het geval van regressie kan dit bijvoorbeeld de absolute waarde van het verschil tussen predictie en observatie zijn, $|\hat{y}_i - y_i|$.
3. Bepaal voor ieder element (x_i, y_i) van D_{kal} op basis van het gekalibreerde model uit stap 1 de non-conformity scores $\{s_i\}_{i=1}^{n_{kal}}$.
4. Bepaal een gewenst betrouwbaarheidsniveau α (bijvoorbeeld 0.05) en bepaal \hat{q} als het $[(n_{cal} + 1)(1 - \alpha)] / n_{cal}$ percentiel van $\{s_i\}_{i=1}^{n_{kal}}$.
5. Bepaal met dit percentiel het betrouwbaarheidsinterval voor de (out-of-sample) validatie set als $C(x_{val}) = \{y: s(x_{val}, y) \leq \hat{q}\}$.

Een betrouwbaarheidsinterval die conform deze procedure is vastgesteld, voldoet aan de zogenaamde 'marginal coverage guarantee', wat inhoudt dat de kans dat de daadwerkelijke observatie binnen het betrouwbaarheidsinterval valt tenminste gelijk is aan $1 - \alpha$ voor iedere non-conformity score en iedere $\alpha \in (0, 1)$. De enige voorwaarde die hiervoor moet gelden, is dat de data onafhankelijk

getrokken is uit dezelfde data, ofwel dat $(x_1, y_1), \dots, (x_N, y_N)$ uitwisselbaar zijn en de volgorde er niet toe doet. Bovenstaande procedure kan uitgebreid worden zodanig dat het resultaat een cumulatieve distributiefunctie is in plaats van een betrouwbaarheidsinterval.

UITBREIDINGEN

De procedure zoals beschreven in de voorgaande sectie is relatief simpel, maar beperkt bruikbaar in de praktijk. Immers, het resulterende betrouwbaarheidsinterval is voor ieder datapunt gelijk. Twee uitbreidingen om deze tekortkoming te adresseren zijn *normalized conformal prediction* en *Mondrian conformal prediction* (zie Boström en Johanson (2022)), die ook met elkaar gecombineerd kunnen worden.

Bij normalized conformal prediction wordt normalisatie toegepast in de score functie:

$$s(x_i, y_i) = \frac{|\hat{y}_i - y_i|}{\sigma_i + \beta}$$

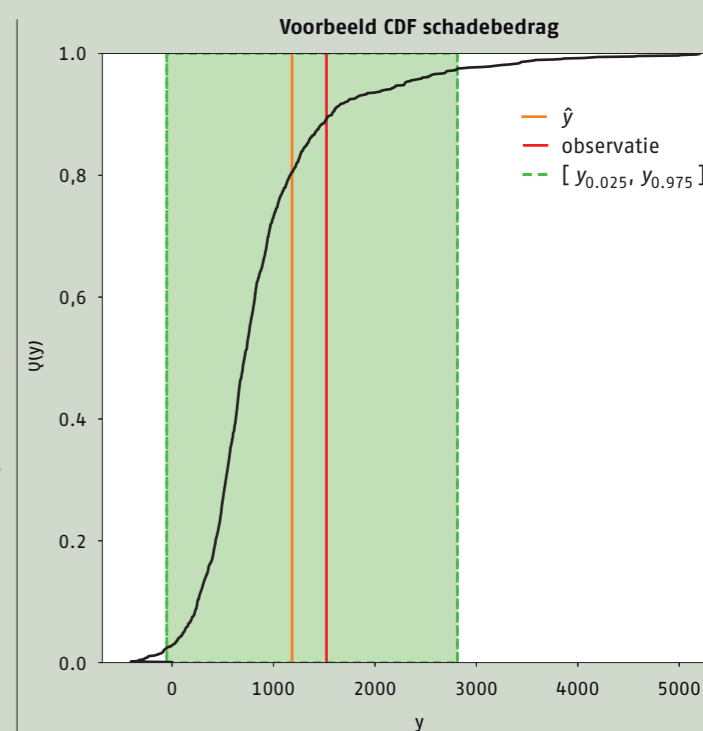
waarbij β een constante is en σ_i een inschatting van de onzekerheid van de schatting \hat{y}_i . Deze parameter kan bijvoorbeeld bepaald worden op basis van een cluster algoritme (zoals K-Nearest Neighbors). Deze variant zorgt er derhalve voor dat het betrouwbaarheidsinterval verschilt voor verschillende datapunten, afhankelijk van σ_i .

Bij Mondrian conformal prediction worden eerst clusters gedefinieerd binnen D_{kal} , en wordt vervolgens de procedure zoals beschreven in de voorgaande sectie toegepast per cluster. Ook dit leidt tot verschillende betrouwbaarheidsintervallen, afhankelijk van het specifieke cluster waarin een datapunt zich bevindt.

Ook bovenstaande technieken hebben als voorwaarde dat de datapunten uitwisselbaar zijn. Echter, voor tijdreeksen is dat niet het geval. Alternatieve conformal prediction procedures zijn hiervoor beschikbaar, zie bijvoorbeeld Xu en Xie (2021).

VOORBEELD

De combinatie van normalized en Mondrian conformal prediction is uitgewerkt voor een voorbeeld data set zoals beschreven in Plat (2017). Deze data komt van de website Kaggle en betreft een autoverzekering portefeuille van de Amerikaanse verzekeraar Allstate. De dataset bevat voor 18.318 observaties 116 categorische en 14 continue (geanonimiseerde) verklarende variabelen en het bijbehorende schadebedrag. Waar in Plat (2017) het doel was om per observatie het schadebedrag zo goed mogelijk in te schatten met het zogenaamde XGBoost algoritme, kan dat nu uitgebreid worden met een betrouwbaarheidsinterval van deze schatting. Figuur 1 laat het resultaat (observatie, schatting, betrouwbaarheidsinterval, cumulatieve verdelingsfunctie) hiervan zien voor één willekeurig datapunt.



Figuur 1: voorbeeld cumulatieve verdelingsfunctie schadebedrag

De figuur laat een redelijk wijd betrouwbaarheidsinterval zien rondom de schatting. De observatie ligt binnen het betrouwbaarheidsinterval, relatief dicht bij de schatting \hat{y} .

CONCLUSIE

In dit artikel is een techniek beschreven waarmee een betrouwbaarheidsinterval of distributie bepaald worden voor ieder mogelijk predictiemodel. Met de aanvulling van deze techniek zijn machine learning modellen meer bruikbaar voor actuariële doeleinden. ■

Referenties

Anastasios, Angelopoulos en Bates (2022), A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification

Boström en Johanson (2022), Mondrian Conformal Regressors

Plat (2017), Data Science en Machine Learning: concreet voorbeeld verzekeringsportefeuille

Xu en Xie (2021), Conformal Prediction Interval for Dynamic Time-Series

Dr. R. Plat AAG RBA is partner van Risk at Work en geeft in die hoedanigheid advies aan verzekeraars, banken en pensioenfondsen op het gebied van waardering en risicomanagement.

