

kadaster



Erwin Folmer 06-11-2024  
Erwin.Folmer@kadaster.nl

# Geo AI toepassing bij het Kadaster

# Data Science @ Kadaster

Multiple teams work on Data Science topics

(Advanced) Spatial Analytics:

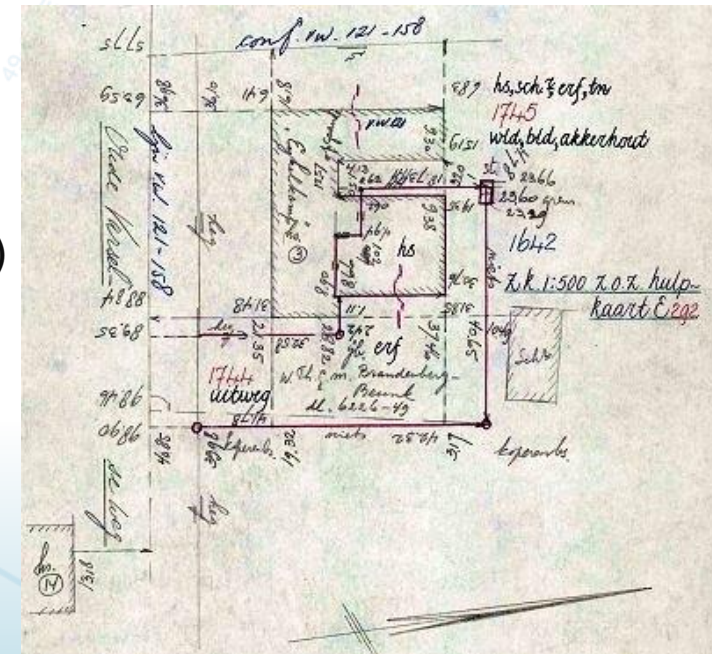
- Customer Insights by Team Research

Focused:

- AI Object Recognition (based on Aerial Photos) (GEC)
- AI Reshaping the Cadastral Map (based on Fieldworks) (KKN)
- AI Optimization Deeds Processing (based on Deeds) (AKTE-AI)

“Other”:

- Data Science Team



Scan de QR codes en ontdek meer over onze projecten!  
Of ga naar [labs.kadaster.nl](https://labs.kadaster.nl)

# 2023 HIGHLIGHTS

## van het DATA SCIENCE TEAM

Het team is in 2019 opgericht om het Kadaster te helpen bij ideeën en vragen rondom Data Science. Hierbij zijn laagdrempeligheid, innovatie en samenwerking belangrijk voor ons.

**Visie 2024**  
Om op toekomstige vragen voorbereid te zijn realiseren we experimentele oplossingen voor veilig data delen en (kunstmatige) data generatie.



### Feitjes

Het team bestaat uit 5 vrouwen en 7 mannen en de gemiddelde leeftijd is 36 jaar. We hebben dit jaar 141 user story's afgerond.



### Voorspelmodel

Het werkaanbod van Rechtszekerheid wordt door een AI model voorspeld. Zo hebben we nu een beter beeld wat er aan aktes binnen gaat komen. Zo'n voorspelmodel ontwikkelen we nu ook voor Landmeetkunde.



### Lock - Unlock

Lock-Unlock wordt uitgevoerd in opdracht van het Federatief Datastelsel (FDS). Binnen het FDS is het belangrijk om op een verantwoorde manier data te kunnen delen. Binnen dit project onderzoeken we de mogelijkheden van een autorisatie ontologie binnen Linked Data en bouwen we verder op de KKG.



### Geodata voor iedereen

Voor het Geodataplein zijn chatbot Loki, de Node Editor en de Query Builder ontwikkeld. Een burger kan vragen stellen aan Loki en een informatie-analist kan met de Query Builder gemakkelijk de KKG bevragen. Voor de developer is de Node Editor ontwikkeld.



### Beelddetectie

Beeldherkenning pasten we toe voor het herkennen van splitsingstekeningen in aktes, objectkaarten in gearchiveerde microfilms en het signaleren van te verwijderen ID-bewijzen in SAP CRM.



### Kadaster Knowledge Graph (KKG)

Met Geonovum is IMX-GEO als gestandaardiseerd model voor de Kadaster Knowledge Graph (KKG) ontwikkeld en deze is geïmplementeerd. De KKG is opnieuw opgebouwd met behulp van ETL pipelines voor eenvoudig beheer.



### Samenwerking Universiteit Twente

Het afgelopen jaar zijn twee Engineering Doctorate (EngD) projecten uitgevoerd: Het voorspellen van graafschades op basis van KLIC-meldingen en het ontwerp van een integrale architectuur voor metadata.



### Hackalod

In de Hackalod 2023 van de Erfgoedsector, hebben we meegewerkt aan de prijswinnende "Koppelaar", waarmee data met elkaar verbonden kan worden.



### Bouw-Woningmarkt-Voortgang dashboard

Voor het ministerie van Binnenlandse Zaken en Koninkrijksrelaties is een dashboard ontwikkeld voor het monitoren van nieuwbouwprojecten per gebied.



### Studenten

In 2023 hebben we 12 studenten begeleid met verschillende onderzoeken. Recreatiewoningen in Zeeland zijn door beelddetectie in kaart gebracht en er is in samenwerking met een IT Trainee een experiment uitgevoerd voor het voorspellen van woningmarkttransacties.



# DATA-DRIVEN PREDICTION AND REDUCTION OF EXCAVATION DAMAGES

Jiarong Li  
EngD candidate  
J.LI-5@UTWENTE.NL



**zoarg**  
graafschade reductie



kadaster

UNIVERSITY  
OF TWENTE.

# BACKGROUND INFORMATION

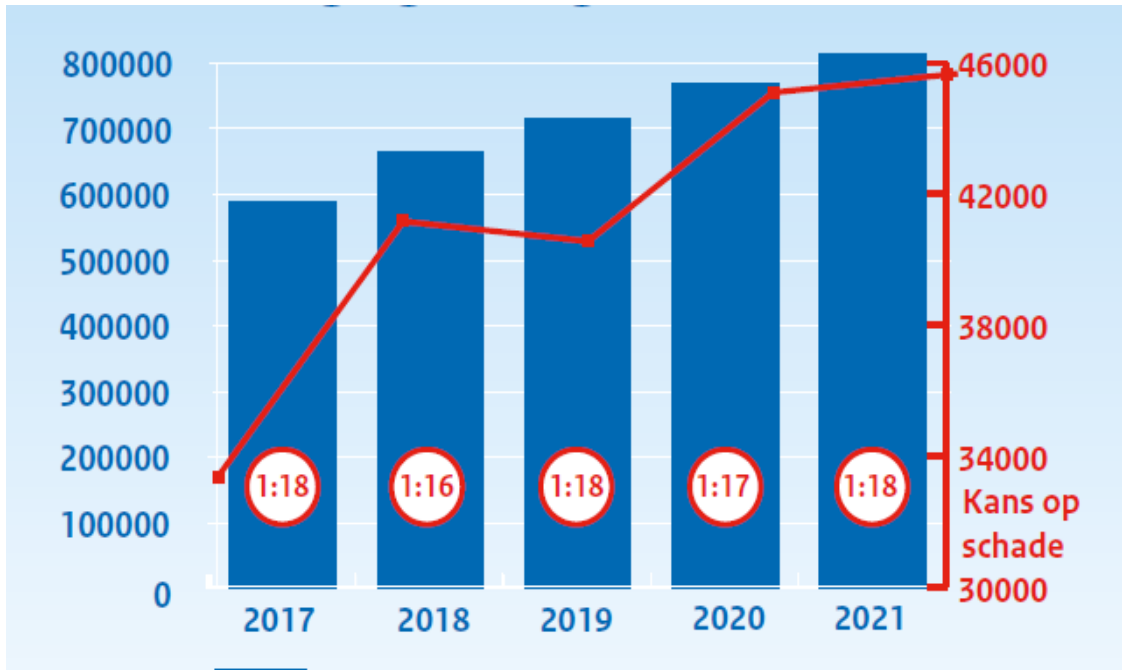
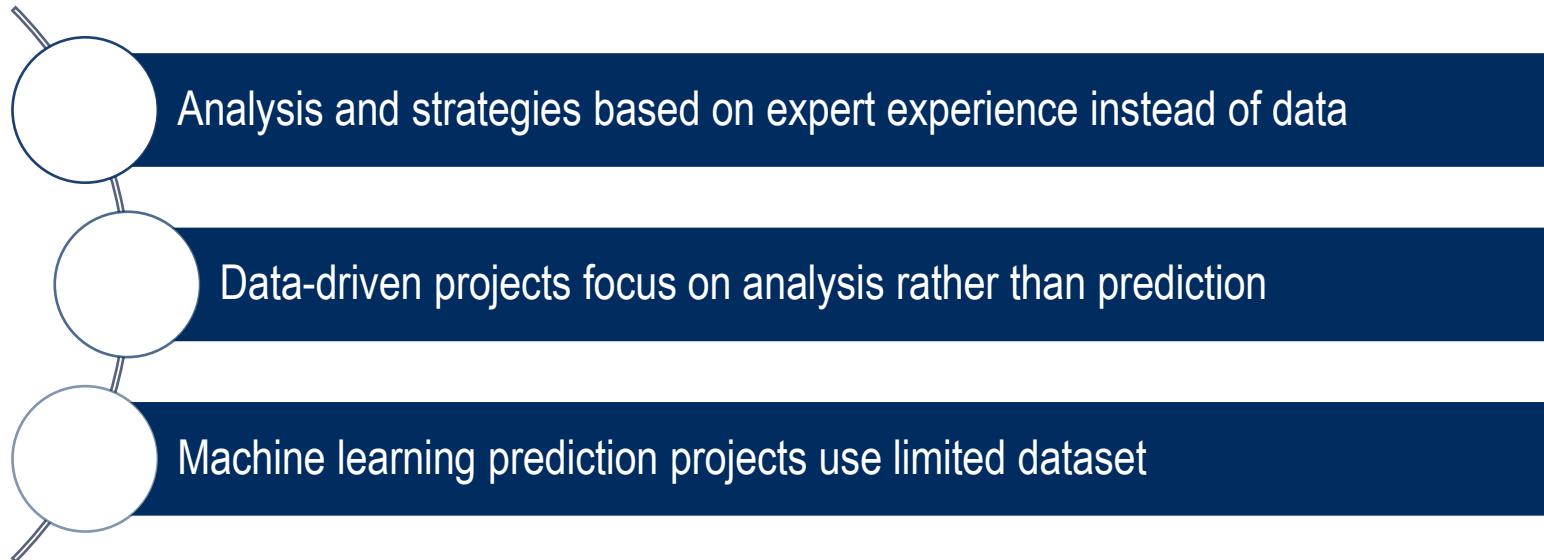


Figure: Number of excavations and excavation damages in the Netherlands from 2017 to 2021

# PROBLEM



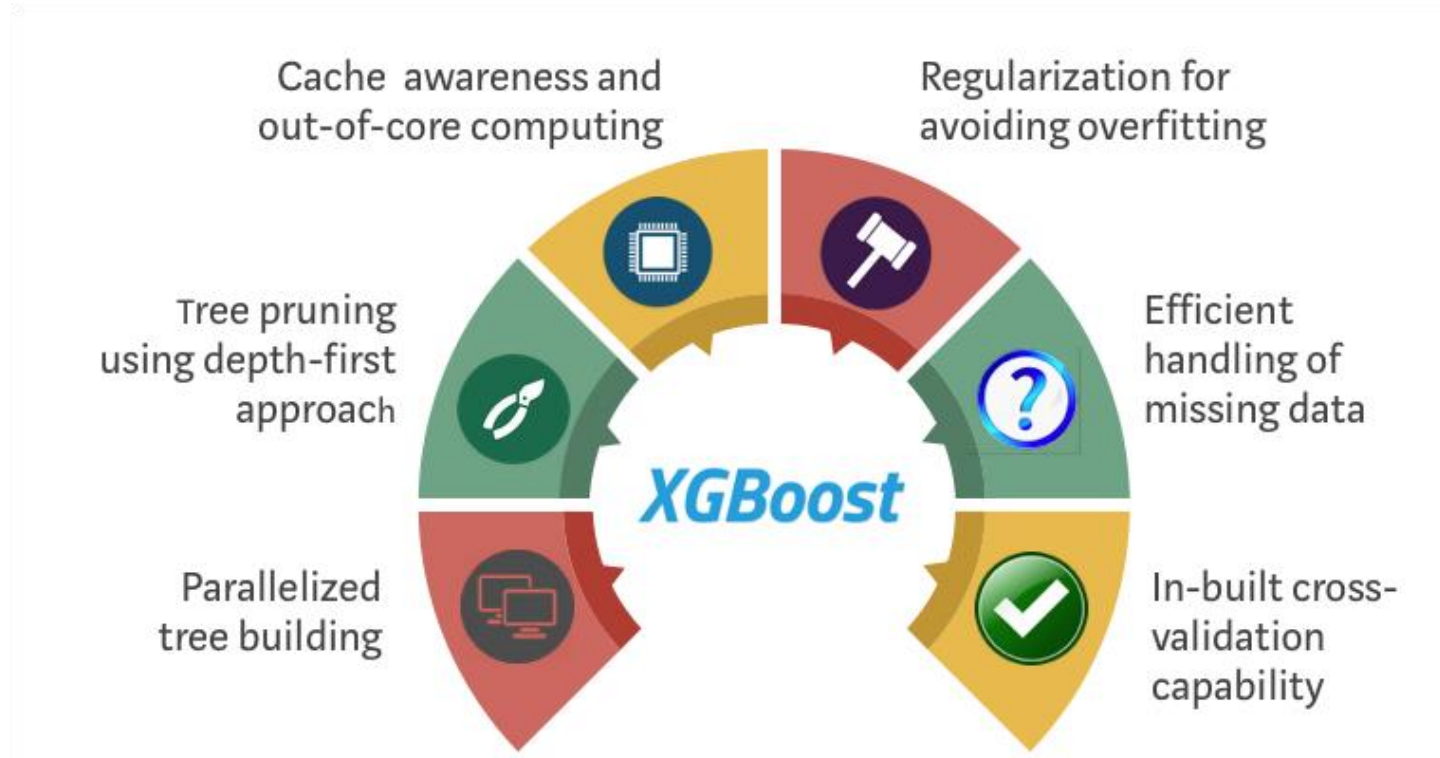
*Data-driven prediction and reduction of excavation damages based on historical data and spatial data*

# DATA

Table 2: Description of datasets for modelling

Dataset name	schademelding	graafmelding	BRO soilmap	Bestand Bodemgebruik	Bomen
Description	excavation damage instances of the Netherlands from 2019 to 2021	excavation instances of the Netherlands from 2019 to 2021	soil type of the Netherlands	landuse of the Netherlands	tree density of the Netherlands
Data format	tabular data(csv)	tabular data (csv)	vector data (shp)	vector data (shp)	raster data (tif)
License	closed data	closed data	open data	open data	open data
Interested columns	-	Gemeente, Oppervlakte, Woonplaats, Provincie, Opdrachtgever, Kliknummer, Oppervlakte, Graafbericht, Prioriteit, Grondroerder, Berichtsoort, EV, Themas, Typenwerkzaamheden, Datum-aanvang-str, Datum-aanvraag-str	soiltype	landuse	tree-density

# PREDICTION SYSTEM - ALGORITHM





# PREDICTION SYSTEM - EVALUATION

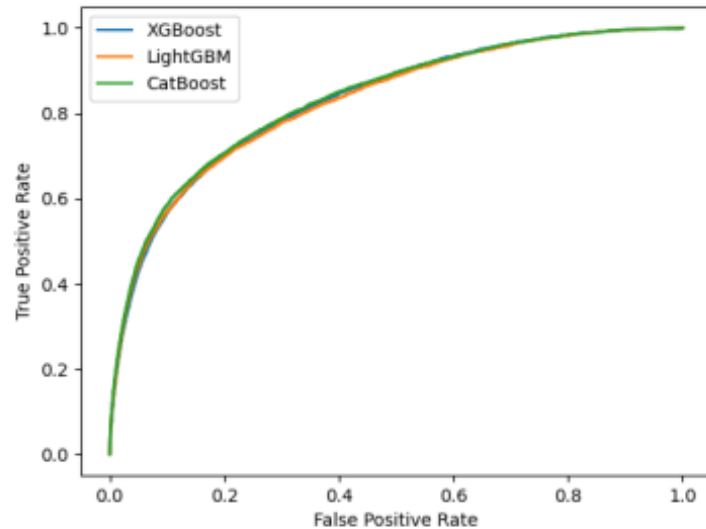


Figure 2: ROC curve of the XGBoost, LightGBM and CatBoost classifiers

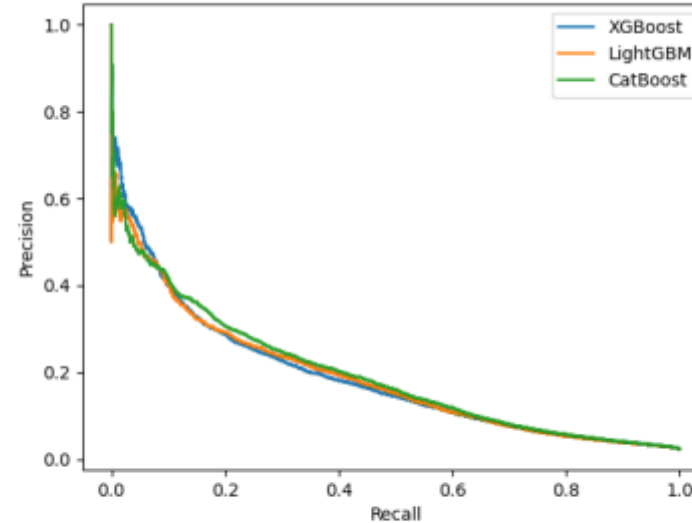


Figure 3: Precision-recall curve of the XGBoost, LightGBM and CatBoost classifiers

**Improved XGBoost model:** AUC: 0.827, Balanced accuracy: 0.747

**Important features:** Tree density, Interval between requesting and working, Polygon complexity, Excavation company and Client company, Landuse (build-up area and highway)

# FEATURE ENGINEERING

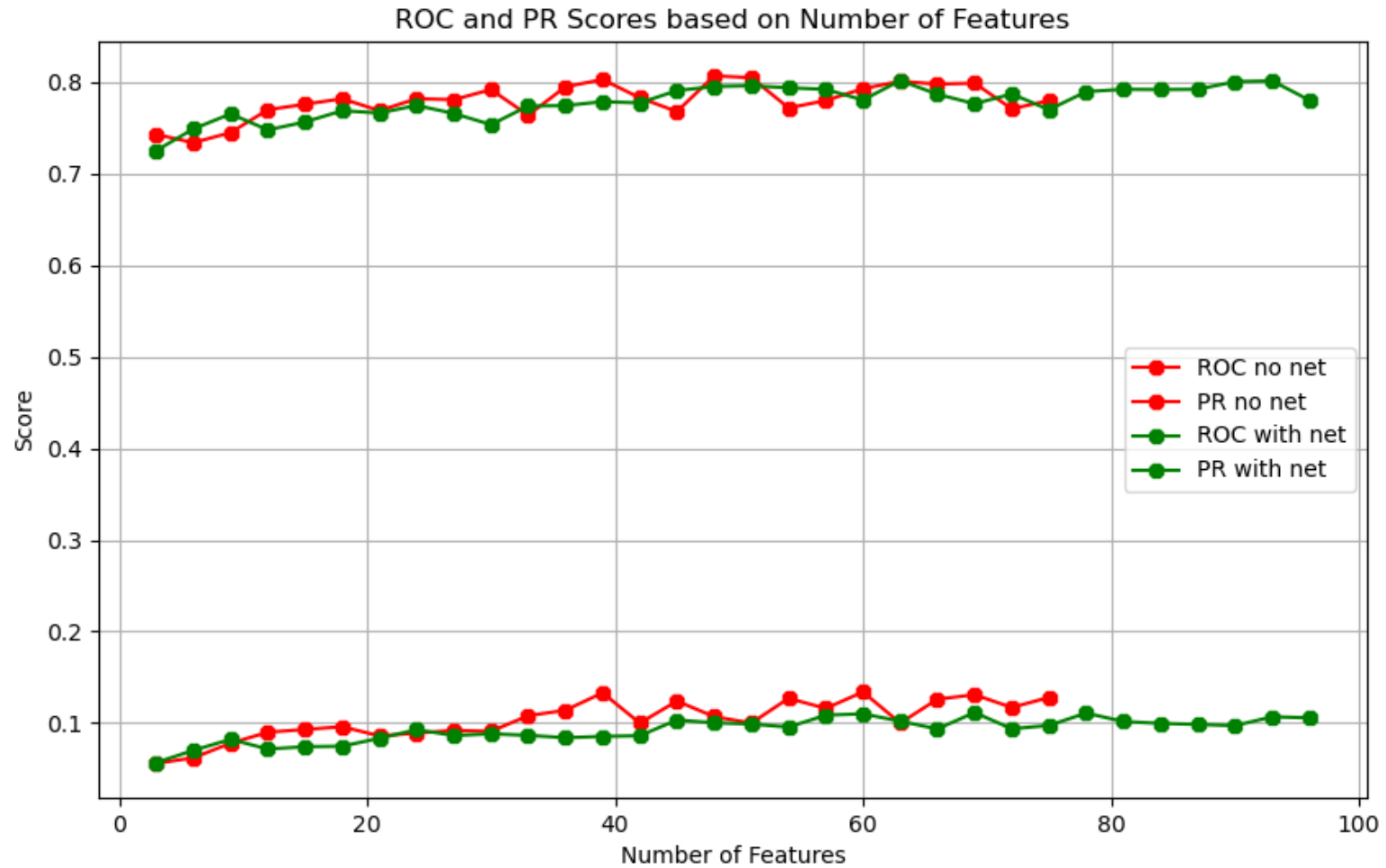
## Objective:

- To test the usefulness of network features in predicting excavation damages
- To find an optimal set of feature combinations for the modelling.

## Dataset:

- Graafmelding and schademelding in 2021 in 3 municipalities.
- Bil-zips from Enexis, KPN and Ziggo in 2021, netinformatie from Brabant-water.
- Landuse, soiltype and tree\_density data.

# FEATURE ENGINEERING



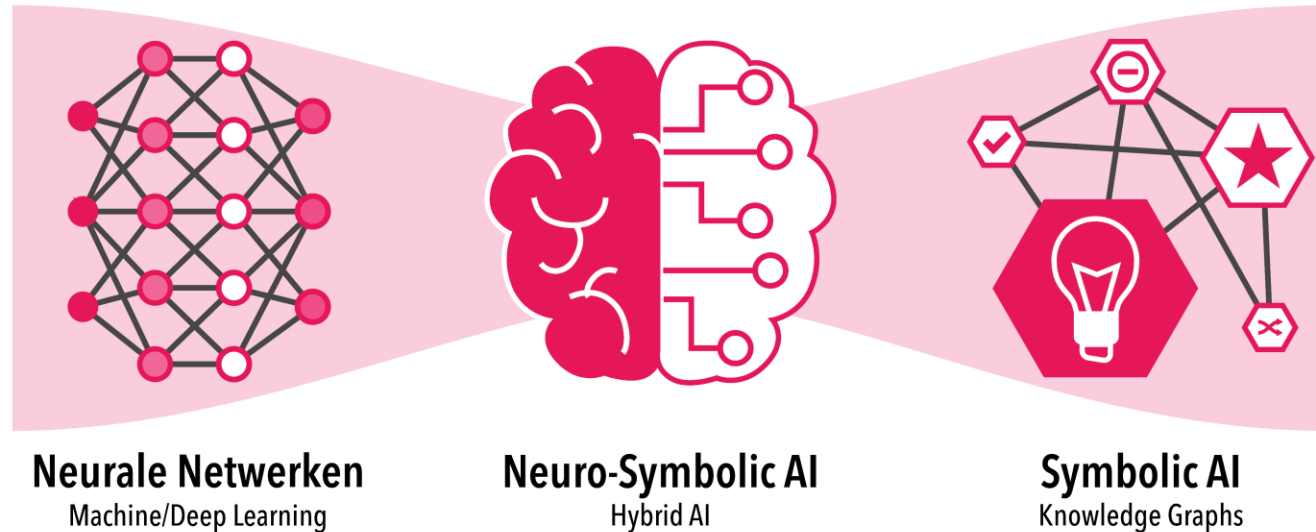
# FEATURE ENGINEERING

## Conclusion

- From ROC the difference is not obvious. From PR the model without utility features is better than the model with utility features.
- No need to add utility features?
  - **Interpretation.**

# CONCLUSIE

- We kunnen graafschade voorspellen
- Drie issues:
  - Voorspellen  $\neq$  **Verminderen** (dus hoe in te zetten?)
  - Draagvlak – Netwerk(geometrie)data (dus hoe hiermee om te gaan?)
  - Data aanpak vs Expertise aanpak (dus: samen?)



# DATA & AI EVENT VAN HET OOSTEN

Lectoraat Applied Data Science & AI Lectoraten Academie IT en Mediadesign

Meld u nu aan!



Op donderdag 21 november is de installatie van Erwin Folmer als lector Applied Data Science en AI tijdens het Data & AI event van het Oosten.

## SNELLE INFO

Datum en tijd	Locatie	Deelname
donderdag 21 november	HAN Campus Arnhem	Deelname is gratis