



Vertraging in de woningbouw voorspellen

Een Shintō Labs-Masterclass

Agenda

- Introductie
- Probleemstelling
- Onderzoek
- Resultaten
- In de praktijk
- Vragen



Introductie

Bouwdoelen verder uit zicht: woningcorporaties bouwen te weinig

CBS: krapte woningmarkt nam vorig jaar verder toe, overbieden is de norm



NOS Nieuws • Donderdag 16 februari, 06:00

Starter heeft kans op slechts 3 procent van de koopwoningen in Nederland



Huizenprijzen dalen verder, maar starters hebben het nog moeilijker

Door RTLZ

10 februari 2023 15:42 • Aangepast 15 februari 2023 14:20



En jonge vader kijkt naar woningen in de etalage

Huizen alsmat schaarser en duurder, en steeds minder verkocht

EIB: meer vertraging voor woningbouw in grote steden dan verwacht

Introductie

Woningmarkt verder in de knel met val kabinet: 'Stilstand is achteruitgang'

Door Suzanne Blotenburg
10 juli 2023 15:03 • Aangepast 10 juli 2023 16:31



Nieuwbouwproject in Gemert. Demissionair minister van Wonen Hugo de Jonge wilde nieuwbouw sneller van de grond laten komen.

Met de val van het kabinet-Rutte IV komen tot ergenis van veel organisaties grote vraagstukken op pauze te staan. Ook de woningmarkt dreigt de dupe te worden van de politieke onrust. Meer nieuwbouw, betaalbare huur? Hoe langer het duurt, hoe meer de boel vastloopt, vrezen kenners.

Onderzoek: het woningtekort houdt nog jaren aan

Huizentekort Volgens onderzoeksbureau ABF Research is de situatie op de Nederlandse woningmarkt het afgelopen jaar verslechterd. Onder meer door de komst van Oekraïense vluchtelingen is het acute woningtekort gestegen.

Milo van Bokkum • 12 juli 2023 om 8:00 • Leestijd 2 minuten



Woningen in aanbouw in het Brabantse Breda.
Foto Robin van Lonkhuijsen/ANP

De komende jaren blijft in Nederland een groot tekort aan woningen bestaan. De voorraad neemt maar beperkt toe, terwijl de bevolking onverwacht snel groeit. Dat blijkt uit onderzoek van het bureau ABF Research, dat jaarlijks voor onder meer provincies, ministeries en brancheverenigingen de situatie op de woningmarkt analyseert.



NOS Nieuws • Vandaag, 07:35 • Aangepast vandaag, 08:17

Woningtekort stijgt fors naar 390.000, pas vanaf 2028 minder krapte

Het tekort aan woningen is in een jaar tijd fors gestegen. Er is in 2023 een tekort van 390.000 woningen. Vorig jaar ging het nog om 315.000 huizen. Pas rond 2028 wordt de krapte op de woningmarkt iets minder.

Introductie

Woningbouw Monitor ≡ Demo omgeving

Projectenkaart

Aantal Woningen (Bruto) **33706**

Aantal Onttrokken **1964**

Totaal Aantal (Netto) **31742**

Woningvoorraad

Beheer Project

Data management

Administratie

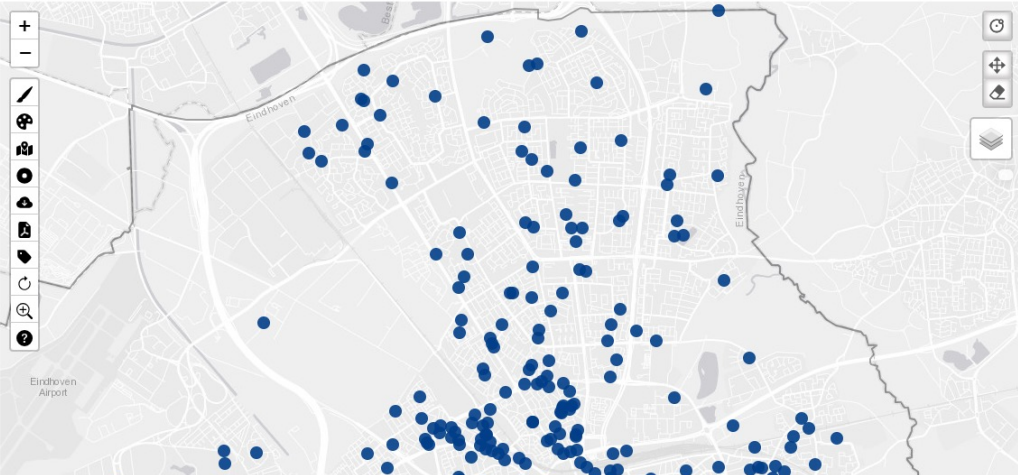
Geschiedenis

Rapportage

Labs

Kaart | Lijst | Rubik's Cube | Overzicht | Prognose | Programma match

Snel overzicht tonen



Actuele data ✓

Reset

Projectnaam ⓘ
Maak selectie ...

Buurt ⓘ
Maak selectie ...

Wijk ⓘ
Maak selectie ...

Aantal woningen in dit project
-150 to 15000

Opleverjaar ⓘ
2023 to 2040

Vertraging in opleverjaar ⓘ

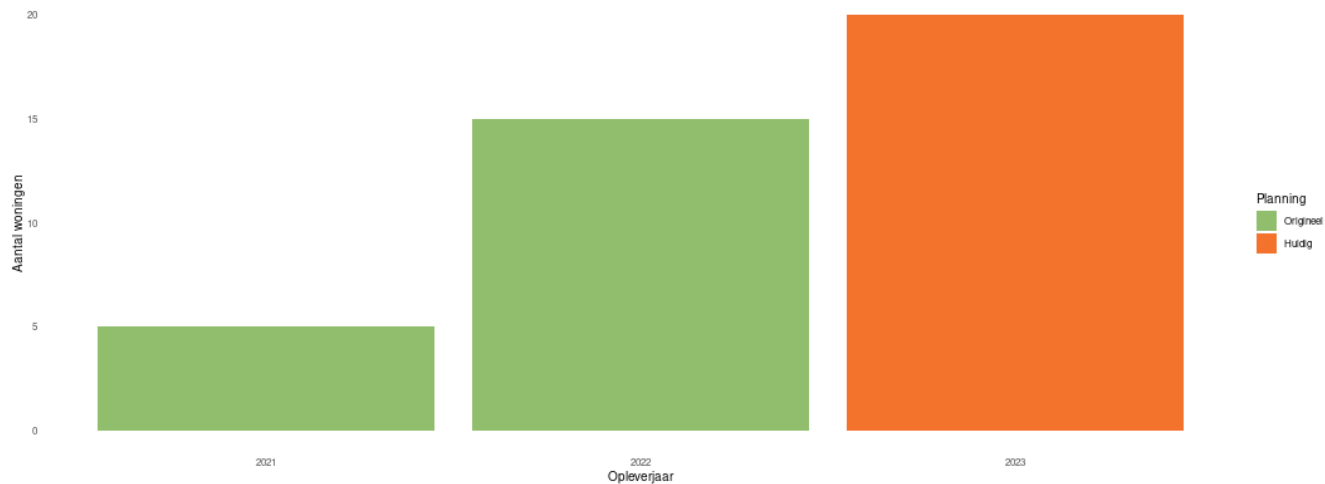
Introductie

20 woningen oplevering in 2023

Info Deelgebieden Locatie **Planning** Geschiedenis

Getoond zijn de 'originele' planning van het project, en de huidige planning na vertragingen in opleverjaar en -maand.

Aanpassingen voor 1 Feb. 2019 worden niet getoond, 'origineel', betekent dus op die datum of later.





Probleemstelling

Probleemstelling

- Voorspellen van vertragingen in de woningbouw met behulp van open data
- Waarom?
- Meer vraag dan aanbod
- Oplossing: aanbod vergroten
- Jaarlijks aantal huizen
- Helpt beleidsmakers doelen behalen
- Eerste schaalbaar onderzoek! (+ open data, bezwaren)

Onderzoeksvragen

Hoofdvraag:

Hoe goed kunnen we vertraging van minstens 1 jaar bij bouwprojecten van woningen voorspellen gebruik makende van classificatie-modellen en open data?

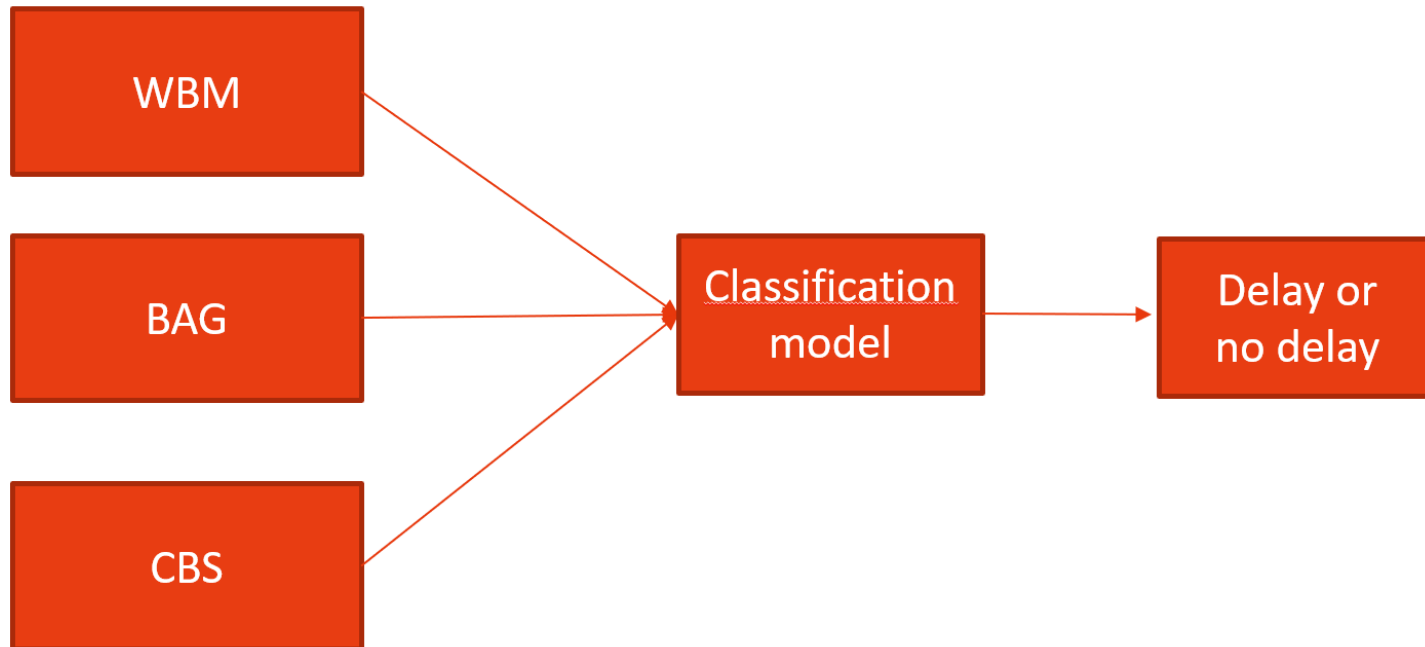
Deelvragen:

- *Welke eigenschappen in open data zijn belangrijk voor de voorspelling van vertraging in bouwprojecten voor huizen?*
- *Welke classificatiemodellen presteren het beste?*



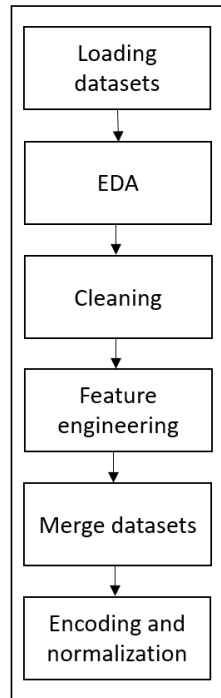
Onderzoek

The how

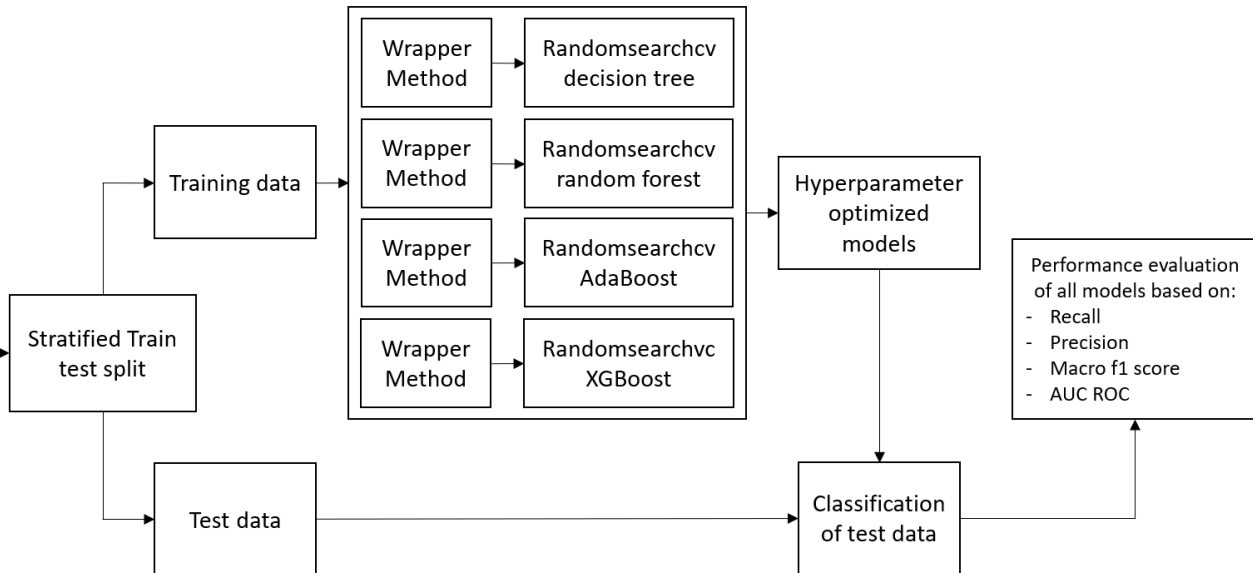


Stappenplan

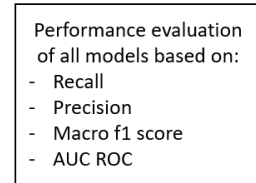
Data load and preprocessing



Feature selection, Hyperparameter optimization and Modeling



Evaluation

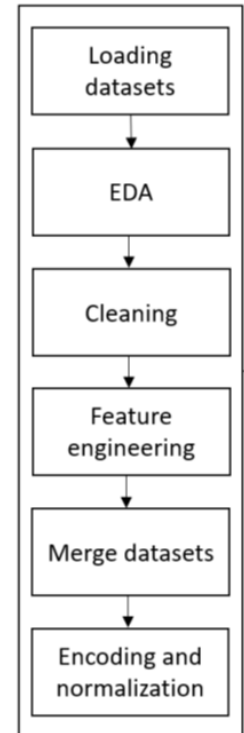


Preprocessing

- Open data focus op demografie (CBS) en omgeving (BAG) bouwproject (WBM)

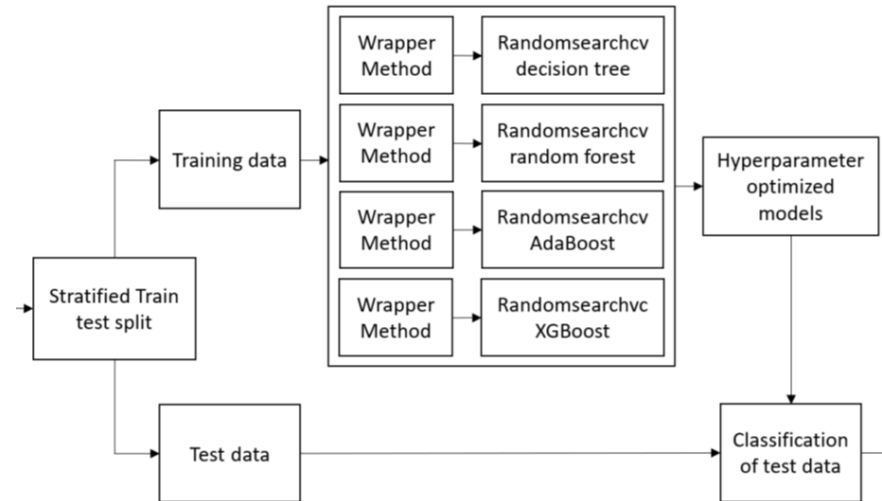
| Dataset | Rows | Columns | Description |
|-----------------------|--------|---------|---|
| WBM of Groningen | 1103 | 26 | Housing construction project data of the municipality Groningen |
| WBM of Leiden | 369 | 30 | Housing construction project data of the municipality Leiden |
| WBM of Eindhoven | 2871 | 26 | Housing construction project data of the municipality Eindhoven |
| BAG extract Groningen | 143488 | 55 | Information about all buildings in the municipality Groningen |
| BAG extract Leiden | 74438 | 55 | Information about all buildings in the municipality Leiden |
| BAG extract Eindhoven | 140372 | 55 | Information about all buildings in the municipality Eindhoven |
| CBS key figures | 373 | 19 | Statistics of the people living in the selected neighborhoods |

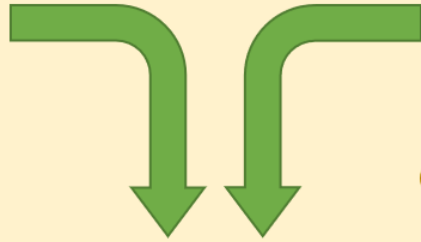
- Cleaning: missing values, data mapping
- Feature engineering: vooral BAG
- Encoding & normalization: helpt de modellen voor betere resultaten
- Merging: 1 dataset met 35 features en label delay/no delay



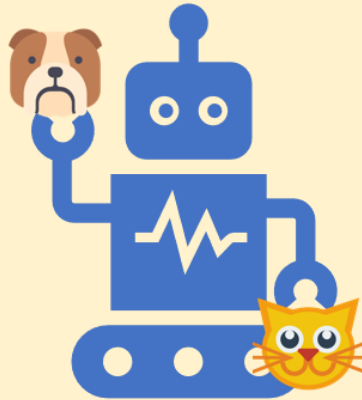
Feature selection en hyperparameter optimization

- Feature selection: wrapper method
- 4 modellen: decision tree, random forest, AdaBoost en XGBoost
- hyperparameter optimization: Random search
- Modeling: train modellen





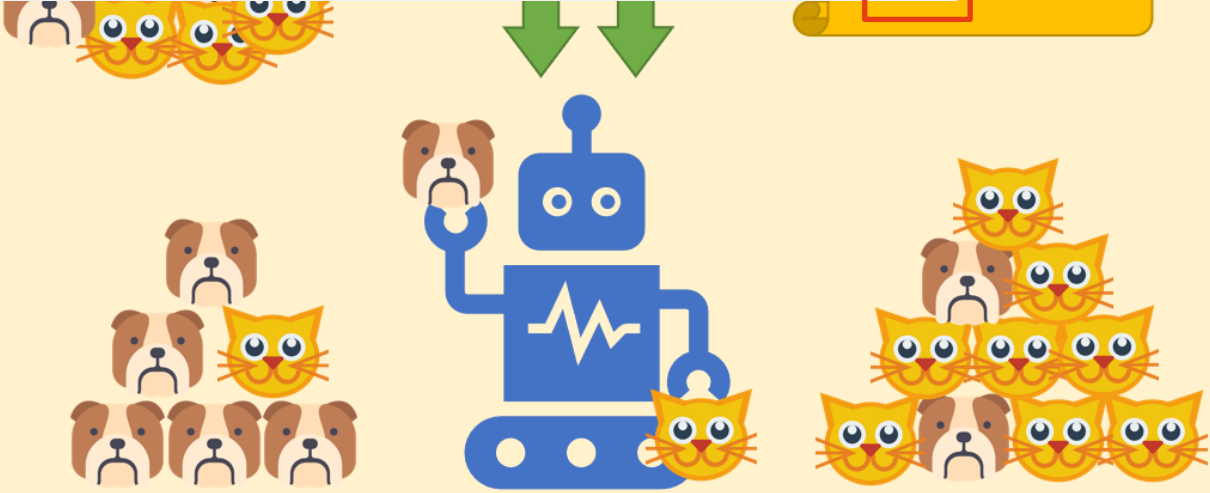
Logistic Regression
SVM
Decision Tree
K Nearest Neighbours
...



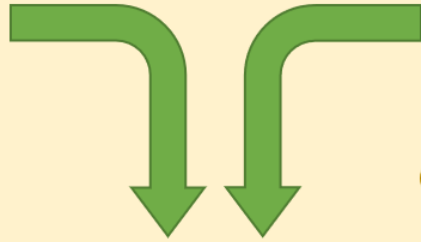
From <https://towardsdatascience.com>

| | fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | alcohol | quality |
|---|---------------|------------------|-------------|----------------|-----------|---------------------|----------------------|---------|------|-----------|---------|---------|
| 0 | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |
| 1 | 7.8 | 0.88 | 0.00 | 2.6 | 0.098 | 25.0 | 67.0 | 0.9968 | 3.20 | 0.68 | 9.8 | 5 |
| 2 | 7.8 | 0.76 | 0.04 | 2.3 | 0.092 | 15.0 | 54.0 | 0.9970 | 3.26 | 0.65 | 9.8 | 5 |
| 3 | 11.2 | 0.28 | 0.56 | 1.9 | 0.075 | 17.0 | 60.0 | 0.9980 | 3.16 | 0.58 | 9.8 | 6 |
| 4 | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |

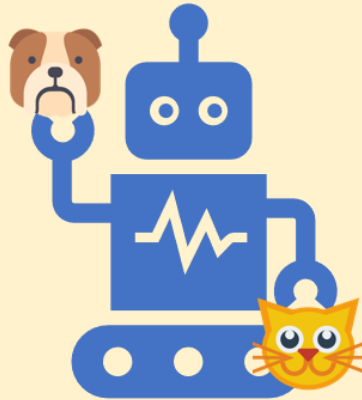
Linear Regression
 SVM
 Decision Tree
 K-Nearest Neighbours
 ...



From <https://towardsdatascience.com>



Logistic Regression
SVM
Decision Tree
K Nearest Neighbours
...



From <https://towardsdatascience.com>



Resultaten

Resultaten

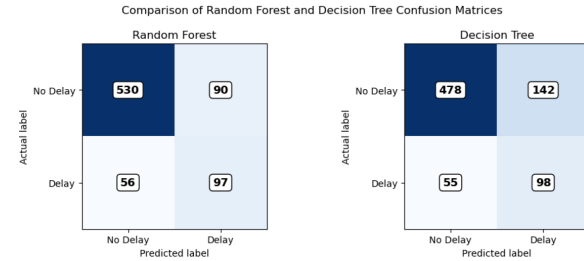
- Recall: Decision tree en random forest

| Metric | Decision Tree | Random Forest | AdaBoost | XGBoost |
|----------------|---------------|---------------|--------------|---------|
| Recall | 0.641 | 0.634 | 0.549 | 0.601 |
| Precision | 0.408 | 0.519 | 0.532 | 0.503 |
| macro f1-score | 0.664 | 0.725 | 0.712 | 0.711 |
| AUC-ROC | 0.74 | 0.81 | 0.78 | 0.80 |

- Feature importance

Belangrijkste features komen uit de WBM:

- Aantal woningen per project
- Prijscategorie van de woningen



Performance evaluation of all models based on:

- Recall
- Precision
- Macro f1 score
- AUC ROC

Classification of test data

Effect open data?

- Resultaten zijn beter in combinatie met open data:

Table 6: Performance metric per classification model using only WBM data

| Metric | Decision Tree | Random Forest | AdaBoost | XGBoost |
|----------------|---------------|---------------|----------|---------|
| Recall | 0.562 | 0.536 | 0.536 | 0.706 |
| Precision | 0.363 | 0.401 | 0.376 | 0.277 |
| macro f1-score | 0.626 | 0.649 | 0.633 | 0.536 |
| AUC-ROC | 0.69 | 0.72 | 0.69 | 0.69 |

- Schaalbaar? Resultaten zijn beter met meerdere gemeentes:

Table 5: Performance metric per classification model using only Eindhoven data

| Metric | Decision Tree | Random Forest | AdaBoost | XGBoost |
|----------------|---------------|---------------|----------|---------|
| Recall | 0.583 | 0.611 | 0.519 | 0.546 |
| Precision | 0.337 | 0.431 | 0.427 | 0.421 |
| macro f1-score | 0.612 | 0.678 | 0.663 | 0.664 |
| AUC-ROC | 0.68 | 0.77 | 0.73 | 0.76 |

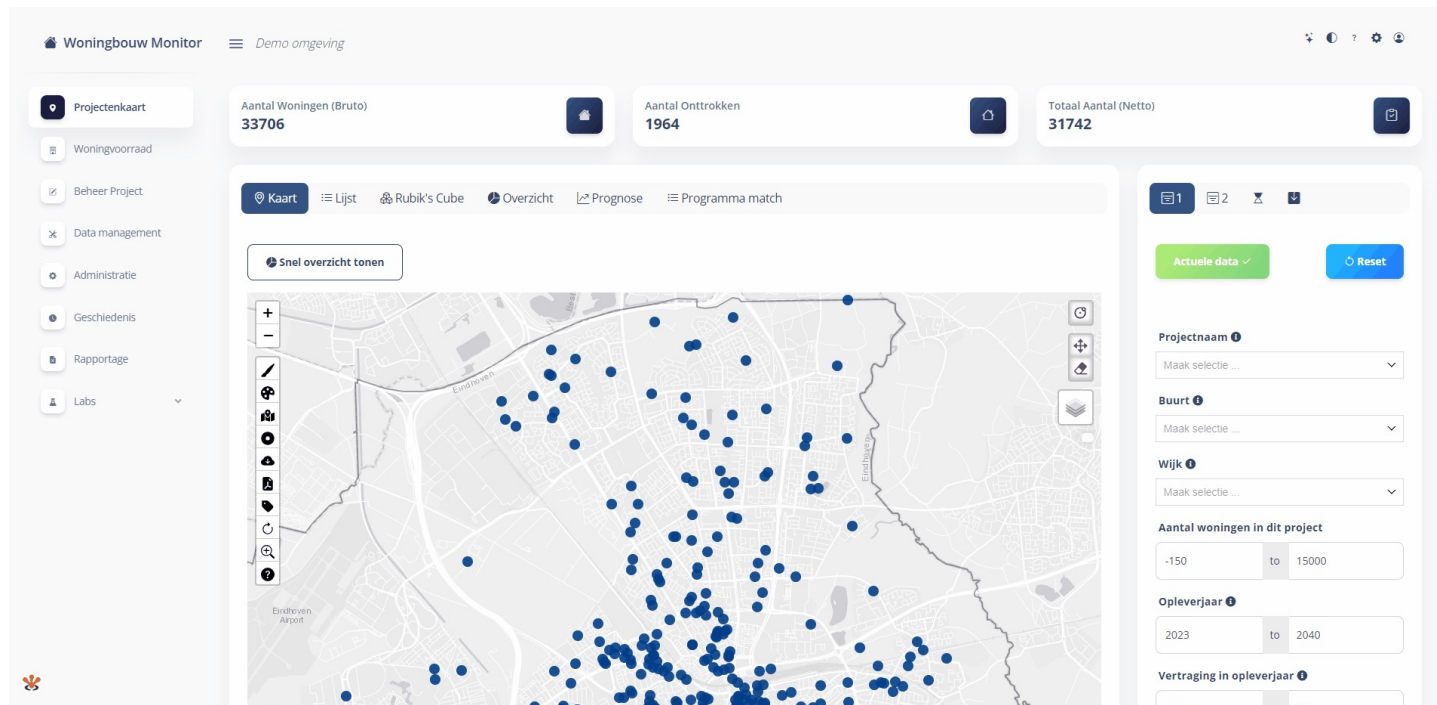
Conclusie

- De resultaten laten zien dat het voorspellen van vertragingen mogelijk is
- Kan als guideline gebruikt worden door de gemeentes
- De methode is schaalbaar en de toevoeging van open data geeft betere resultaten



In de praktijk

In de Woningbouwmonitor



In de Woningbouwmonitor

Projectomschrijving

✍️ [Wijzig project omschrijving](#)

🕒 Audit

🕒 [Audit openen ...](#)

[Vertragingen beheren ...](#)

📷 Afbeelding

📷 [Uploaden](#)



🕒 Voorspelling

Op basis van de instellingen van dit project is er een X% kans op vertraging van Y jaar?

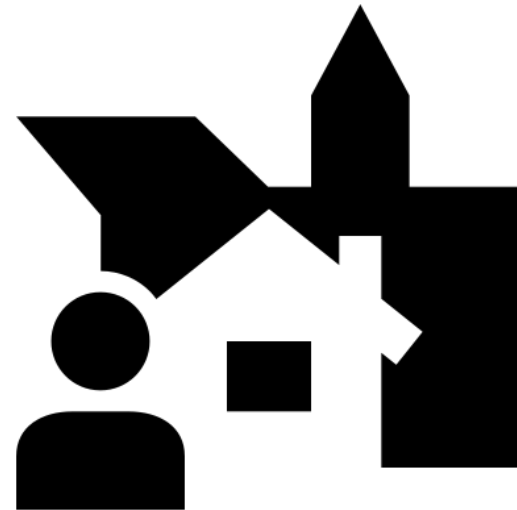
Voorspelling gebaseerd op:

- Aantal woningen die gebouwd worden
-

[→ Lees meer](#)

Voor de overheid

- Snel inzicht waar risico is op vertraging
- Beter beleid op risicovolle projecten
 - Samenkomsten met medebewoners
- Op gemeentelijk, regionaal en provinciaal niveau inzicht
 - Woondeals
 - Doelstellingen halen
- Mogelijk nieuwe inzichten voor het nieuwe kabinet!



Voor de maatschappij

- Verkleinen van het woningtekort
- Inzichten voor burgers
 - Kan ik straks ergens wel een huiskopen?
 - Hoe staat het met “mijn” bouwproject?
 - Hoe lang gaat het bouwen in mijn buurt nog duren?
 - Hoe ziet mijn buurt eruit over x jaar?
- Minder druk op de woningmarkt voor onze burgers





Vragen



Dank voor uw aandacht!
www.shintolabs.nl