

Adreskwaliteit: persoons- en adreskenmerken van onderzochte adressen in de gemeente Groningen tussen 1997 en 2020

Doel onderzoek

De adressituatie (de combinatie van persoon en adres) is het meest gevoelig voor fouten in de Basisregistratie Personen (BRP). De werkwijze van Burgerzaken omtrent foute registraties is reactief en vraagt veel capaciteit: alle signalen worden onderzocht en leiden, indien nodig, tot een aanpassing in de BRP. Het doel van dit onderzoek is na te gaan of er, aan de hand van patroonherkenning, een risicoprofiel van foutieve adressituaties kan worden samengesteld op basis waarvan kan worden geschat of signalen verder moeten worden onderzocht. Dit kan helpen de medewerkers van Burgerzaken te ontlasten.

Data & methode



75.187 adresonderzoeken waarvan ruim 51.000 (68%) aanpassingen ● en ruim 24.000 (32%) niet-aanpassingen ▲

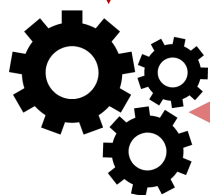


kenmerken
persoon



kenmerken
adres

gelabelde
traindata



univariate analyse
bivariate analyse
logistische regressie
polynomial SVM
decision tree
random forest
gradient boosting

ongelabelde
testdata

classificatie

Resultaten



- jaren woonachtig in Groningen
- ingeschreven in buitenland



- geen



- kenmerken: verschil met voorgaand onderzoek
- decision tree minst goede prestatie
- random forest en gradient boosting beste prestatie
- minder accuraat dan 'medewerker Burgerzaken'



- meer data
- aanscherping aanpassing
- aanscherping onderzoekspopulatie
- doorontwikkelen methoden

Algoritme	Accuracy	Precision	Recall	Sensitivity	ROC
Logistische regressie	63%	80%	61%	68%	.69
SVM	63%	81%	60%	69%	.65
Decision Tree	60%	82%	53%	75%	.64
Random Forest (Random Search)	64%	82%	61%	71%	.64
Gradient Boosting, Model I	64%	83%	59%	74%	.66

Loes van Rijsewijk
08-07-2020
Eindonderzoek ITANN



onderzoek
informatie
statistiek
 groningen

