Translationele gegevenswetenschap

Natuurlijke taalverwerking in de geestelijke gezondheid

Spreker: *Prof. Dr. Marco Spruit (LUMC/LIACS)*

Missie: *"Vertalen van AI technologieën naar praktisch bruikbare oplossingen in populatiegerichte zorg en welzijn"*

Agenda

De ontwikkelingen op het gebied van kunstmatige intelligentie gaan sneller dan ooit tevoren. Wie kent niet ChatGPT en wie staat er niet versteld van de kwaliteit van deze AI? In deze presentatie bespreek ik hoe we nieuwe data -en AI-technologieën kunnen inzetten om uiteindelijk de zorgpraktijk te verbeteren. Dit is het domein van de translationele datawetenschap. Een concrete casus illustreert deze benadering, waarin we klinische teksten analyseerden om het risico op agressie-incidenten te voorspellen bij psychiatrische patiënten.

Health Campus Den Haag

LUMC

LEIDEN BIO SCIENCE PARK

liacs

Over… Marco Spruit

Pretpakket VWO
Propedeuse Nederlands
Propedeuse Muziekwetenschap
Bovenbouwstudie α-Informatica

Als Software Ontwikkelaar

1993

- *Information Retrieval* programmeur
  - ZyLAB Europe BV

1995

- *Big Data* systeemontwikkelaar
  - Koninklijke Marine

1997

- Productsoftware ontwikkelaar/ondernemer
  - Insertable Objects, Wizzer BV

Als Wetenschappelijk Onderzoeker

2003

- OiO in Computationele Linguïstiek
  - Universiteit van Amsterdam (UvA)

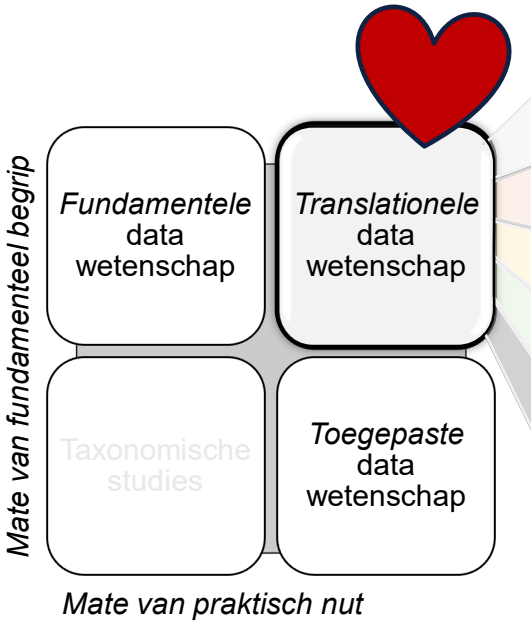
2007

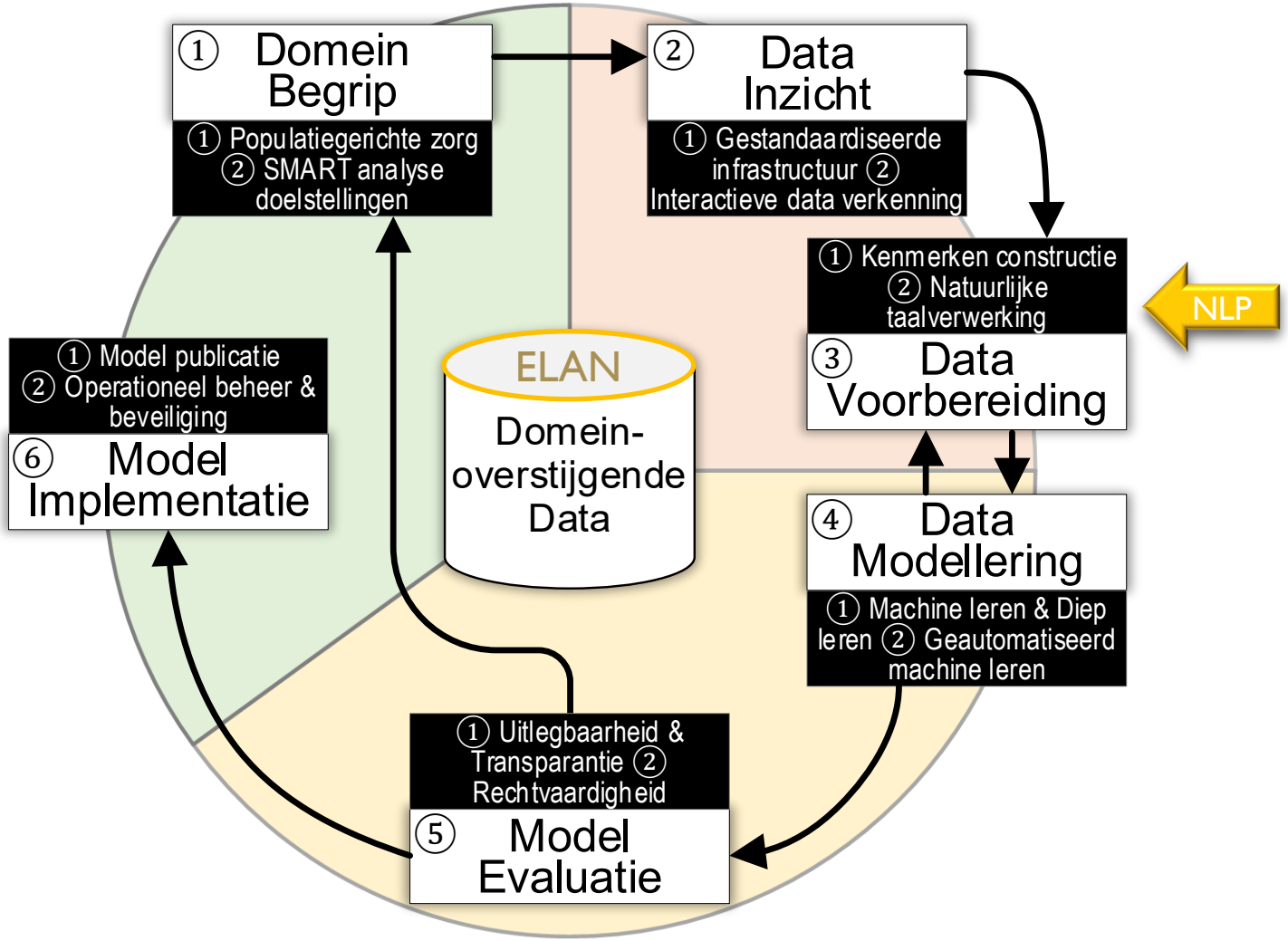
- Universitair docent/hoofddocent Informatiekunde
  - Universiteit Utrecht >> *Applied Data Science* Lab (2015)

2020

- **Hoogleraar *Advanced Data Science in Population Health***
  - **LUMC/LIACS** @ Universiteit Leiden, Health Campus Den Haag
  - PH Living Lab, CAIRELab, TDS Lab, SIG Health Data Science

Translationele Datawetenschap in Populatiegerichte Zorg





Data Techniek Data Analyse e-Health Implementatie

Spruit, Marco. (2022). Translational Data Science in Population Health (p. 20). Leiden University. https://doi.org/10.5281/zenodo.7665858

Casus in vogelvlucht

Voorspellen van agressie risico bij opgenomen psychiatrische patiënten mbv klinische notities in het EPD

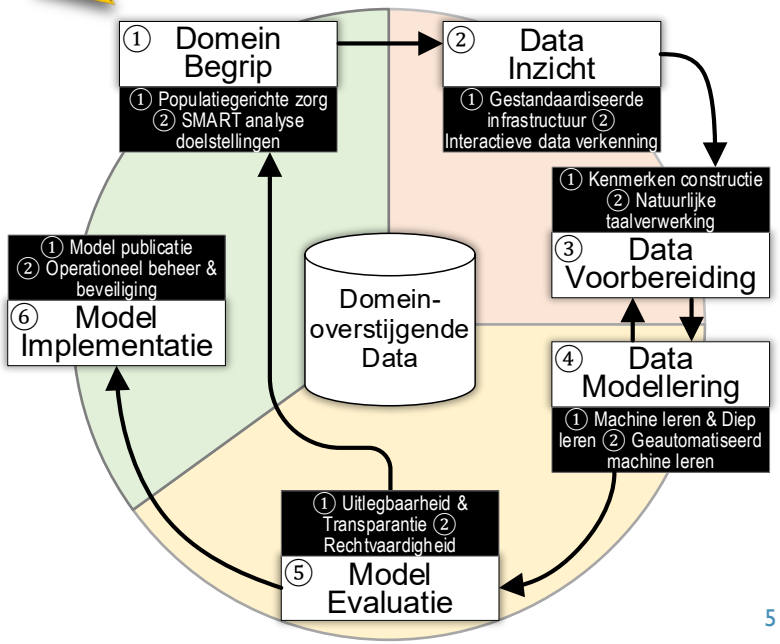
Korte casus: Artificiële intelligentie (NLP & ML) in translationele datawetenschap via CRISP-DM

Menger,V., Spruit,M., Est,R. van, Nap,E., & Scheepers,F. (2019). Machine Learning Approach to Inpatient Violence Risk Assessment Using Routinely Collected Clinical Notes in Electronic Health Records. *JAMA Network Open, 2*(7), e196709. [online]

JAMA Network | Open

FASE 1/6:
DOMEIN BEGRIP: HET DOEL FORMULEREN

- "Predict for which admissions a violence incident will occur in the first 30 days, based on clinical texts that are written up to and including the first day of admission"
  - Prediction task excludes incidents on Day 1 of admission
    - insufficient data available to make a prediction
  - 30 days interval chosen for sufficient specificity
    - majority of incidents included
    - mean duration of admission is 40.3 days
    - 81.9% of incidents happen during the first 30 days
- Area Under Curve (AUC) to report performance



Data Techniek Data Analyse e-Health Implementatie

Fase 2/6: Data begrip → Beschrijf!

- *Site 1:* UMC Utrecht
- *Site 2:* Antes, Parnassia Group, Rotterdam

Table 1. Descriptive Statistics of the Data Sets Obtained From the 2 Sites

Characteristic	No. (%) Site 1	No. (%) Site 2
Demographic characteristics		
Age, mean (SD), y	34.0 (16.6)	45.9 (16.6)
Men	1536 (48.2)	2097 (64.5)
Data set		
Admissions, No.	3189	3253
Unique patients, No.	2209	1919
Length of stay, median (IQR), d	16.0 (6.0-41.0)	15.0 (5.0-40.5)
No. of words in notes, median (IQR)	2091 (1541-2981)	1961 (1160-3060)
Admissions with violent incidents	290 (9.1)	247 (7.7)
Incidents		
During admission, No.	962	652
During first 4 wk	658 (68.4)	318 (48.8)
During first 24 h	90 (9.4)	42 (6.4)
Staff Observation Aggression Scale-Revised score, median (IQR) [range]	12.0 (8.0-16.0) [2-21]	11.0 (7.0-14.0) [2-19]
Diagnostic and Statistical Manual		

Fase 2/6:
Data begrip → Bekijk de bron!

(2012-07-29)
"Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat de klachten die zij gisteren aan haar voeten ervaarde verdwenen zijn. Mw ging na 4.00 uur weer naar bed en kwam niet meer uit haar kamer tot de ochtend."

?

Fase 2/6:
Data begrip → Voorspel zelf!

(2012-07-29)
"Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat de klachten die zij gisteren aan haar voeten ervaarde verdwenen zijn. Mw ging na 4.00 uur weer naar bed en kwam niet meer uit haar kamer tot de ochtend."

Fase 3/6: Data preparatie

Text representation

- Represent all clinical notes related to 1 admission as 1 vector (not words)

- *paragraph2vec*
- *SVM classifier*

(2012-07-29)
"Mw heeft matig geslapen, sliep van 1.00 uur tot 4.00 uur. Kwam toen uit bed, at koekjes en dronk thee. Nog geadviseerd medicatie te nemen en mijn zorgen geuit over evt. doorschieten in een manie. Mw was er niet gevoelig voor en reageerde geagiteerd. Mw had spreekdrang maar gaf aan dat wanneer zij zich goed voelt ook veel praat. Mw gaat vandaag naar <PERSOON-1> met haar zoon, ziet daar nu niet meer tegenop omdat

[0.341, -0.359, 0.7, 0.926, -0.004, …, -0.129]

weer naar bed en kwa niet meer uit haar kamer tot de ochtend.

(2012-08-05)

[Positive, Negative]

van mijn been" […]

VOORBEELD…

Place your text below:

De ontwikkelingen op het gebied van kunstmatige intelligentie gaan sneller dan ooit tevoren. Wie kent niet ChatGPT en wie staat er niet versteld van de kwaliteit van deze AI? In deze presentatie bespreek ik hoe we nieuwe data -en AI-technologieën kunnen inzetten om uiteindelijk de zorgpraktijk te verbeteren. Dit is het domein van de translationele datawetenschap. Een concrete casus illustreert deze benadering, waarin we klinische teksten analyseerden om het risico op agressie-incidenten te voorspellen bij psychiatrische patiënten.

Specify model and parameters to generate dataset:

Model: CBOW

The Word2Vec (CBOW) model trains multiple words (in the form of bag-of-words) surrounding a target word.

Window size: 3

☐ Negative sampling?

Generate dataset

Tokens:

No.	Token	Freq
1	de	4
2	ontwikkelingen	1
3	op	2
4	het	3
5	gebied	1

Vocabulary size: 61

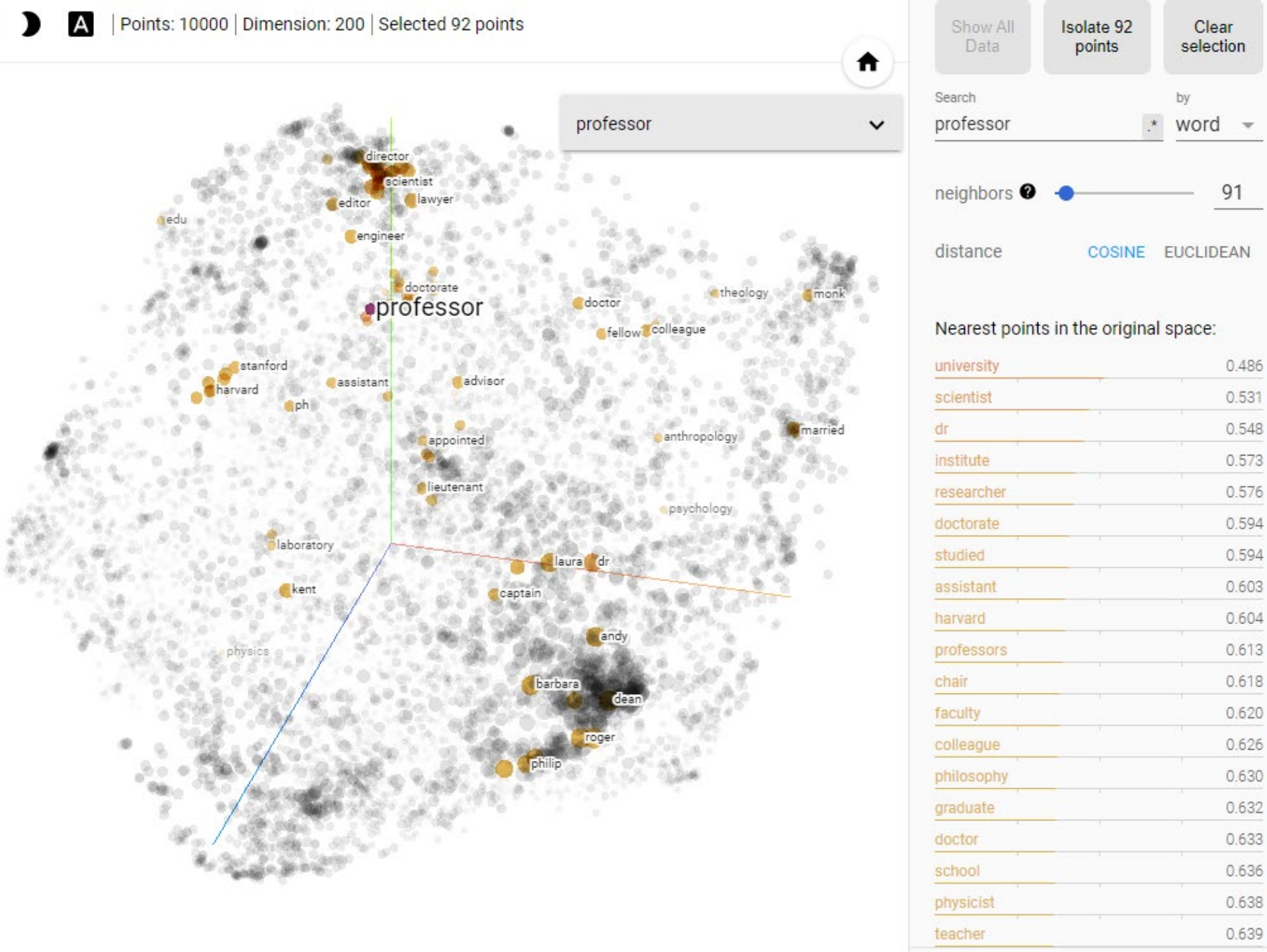
Training examples:

3	de ontwikkelingen het gebied van	op
4	de ontwikkelingen op gebied van kunstmatige	het
5	ontwikkelingen op het van kunstmatige intelligentie	gebied
6	op het gebied kunstmatige intelligentie gaan	van
7	het gebied van intelligentie gaan sneller	kunstmatige

No. of training examples: 77

https://remykarem.github.io/word2vec-demo/

VISUALISATIE.



https://projector.tensorflow.org/

Fase 3/6:
Data preparatie → Vindt het beste model

Menger,V., Scheepers,F., & Spruit,M. (2018). Comparing Deep Learning and Classical Machine Learning Approaches for Predicting Inpatient Violence Incidents from Clinical Text. *Applied Sciences, 8*(6), Data Analytics in Smart Healthcare, 981. [JIF: 2.679]pdf] [online]

- In previous work, we determined SVM as an appropriate classifier for VRA, based on literature and experiments

Table 4. The performance for optimal hyperparameter values for each of the representations combined with the models, based on a 5-fold stratified cross validation. The performance is measured in AUC, along with its standard deviation. The best performance over different models is marked with an [a], the best performance over representations with a [b].

Model	Bag-of-Words Binary	Bag-of-Words tf-idf	Word Embeddings	Document Embeddings
RNN [1]	0.771 ± 0.018 [b]	0.753 ± 0.031	0.654 ± 0.043	0.788 ± 0.018 [a,b]
CNN [2]	0.729 ± 0.030	0.716 ± 0.038	0.684 ± 0.038	0.763 ± 0.024 [a]
NN [3]	0.727 ± 0.033	0.717 ± 0.038	0.751 ± 0.036 [a]	0.745 ± 0.022
NB [4]	0.686 ± 0.026	0.704 ± 0.034 [a]	0.700 ± 0.051	0.692 ± 0.046
SVM [5]	0.759 ± 0.040	0.756 ± 0.036 [b]	0.764 ± 0.024 [b]	0.770 ± 0.029 [a]
DT [6]	0.727 ± 0.018 [a]	0.719 ± 0.041	0.685 ± 0.041	0.665 ± 0.035

[1] Recurrent Neural Network; [2] Convolutional Neural Network; [3] Neural Network; [4] Naive Bayes; [5] Support Vector Machine; [6] Decision Tree.

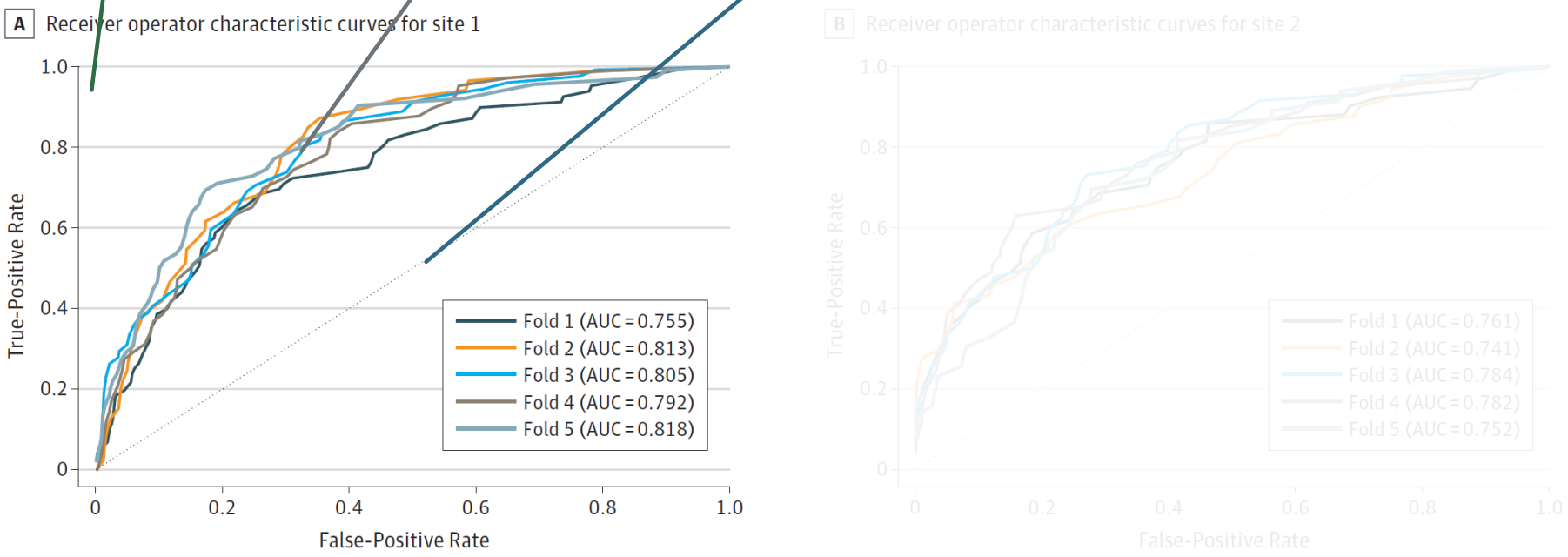
FASE 4/6: (SVM ALS CLASSIFIER)
DATA MODELLEREN → EVALUEER DE PRESTATIE VAN HET VOORSPELMODEL

Perfect Score

Actual Score

Random Score

Figure. Receiver Operator Characteristic Curves for Internal Cross-validations



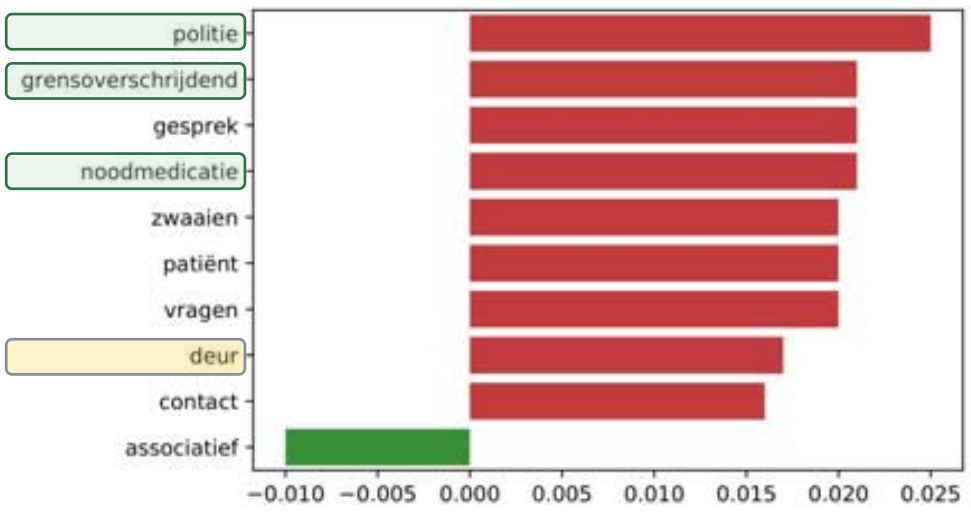
Receiver operator characteristic curves are shown for each fold, according to internal cross-validation in site 1 (A) and site 2 (B). Dashed diagonal lines denote an area under the curve (AUC) of 0.5, ie, predictive validity equivalent to chance. AUC indicates area under the curve.

Fase 5/6:
Model evaluatie → Bespreek met domein experts de data-analyse

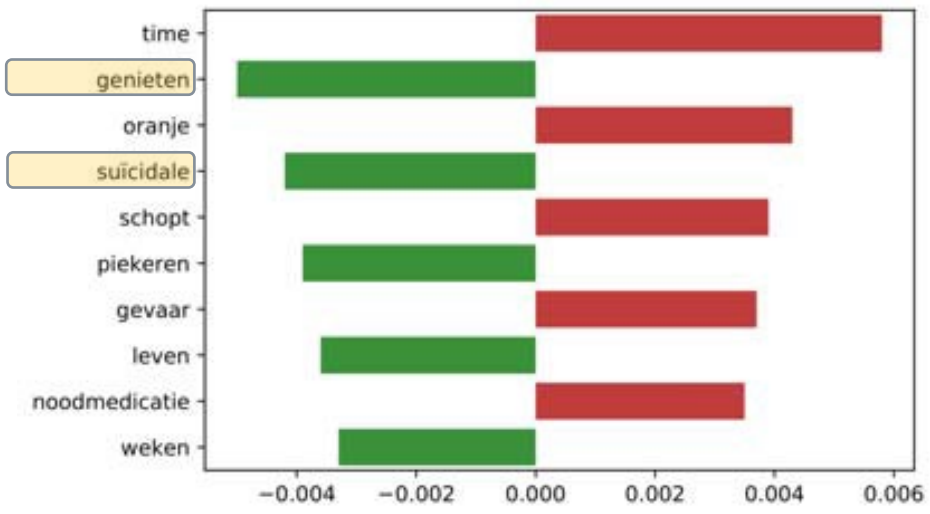
Table 3. Results of Exploratory Analysis

	Site 1				Site 2			
Rank[a]	Term (English Translation)[b]	Ratio	MCC (95% CI)[c]	*P* Value[d]	Term (English Translation)[b]	Ratio	MCC (95% CI)[c]	*P* Value[d]
1	Agressief (aggressive)	1.00	0.17 (0.13 to 0.21)	<.001	Verbaal (verbal)	1.00	0.14 (0.10 to 0.18)	<.001
2	Reageert (reacts)	1.00	0.15 (0.11 to 0.19)	<.001	Dreigend (threatening)	1.00	0.13 (0.08 to 0.16)	<.001
3	Aangeboden (offered)	1.00	0.14 (0.11 to 0.18)	<.001	Agressie (aggression)	1.00	0.15 (0.11 to 0.17)	<.001
4	Boos (angry)	1.00	0.16 (0.12 to 0.19)	<.001	Hierop ([up]on this)	1.00	0.13 (0.09 to 0.16)	<.001
5	Deur (door)	1.00	0.14 (0.10 to 0.18)	<.001	Kantoor (office)	1.00	0.12 (0.08 to 0.16)	<.001
6	Loopt (walks)	1.00	0.15 (0.11 to 0.18)	<.001	Personeel (staff)	1.00	0.12 (0.07 to 0.16)	<.001
7	Ibs (arrest)	1.00	0.14 (0.10 to 0.17)	<.001	Aangesproken (spoke to)	1.00	0.11 (0.08 to 0.15)	<.001
8	Aanbieden (offer)	1.00	0.12 (0.08 to 0.15)	<.001	Agressief (aggressive)	0.99	0.11 (0.08 to 0.15)	<.001
9	Noodmedicatie (emergency medication)	0.99	0.14 (0.10 to 0.17)	<.001	Gevaar agressie (danger aggression)	0.99	0.11 (0.07 to 0.15)	<.001
10	Liep (walked)	0.99	0.12 (0.08 to 0.16)	<.001	Agitatie (agitation)	0.99	0.11 (0.07 to 0.14)	<.001
11	Agressie (aggression)	0.99	0.13 (0.09 to 0.18)	<.001	Geirriteerd (irritated)	0.99	0.10 (0.06 to 0.14)	.001
12	Vraagt (asks)	0.99	0.13 (0.10 to 0.17)	<.001	Separeer (seclusion room)	0.99	0.10 (0.06 to 0.15)	<.001
13	Status vrijwillig (status voluntary)	0.99	−0.12 (−0.14 to −0.09)	<.001	Loopt (walks)	0.99	0.11 (0.08 to 0.14)	.02
14	Psychotisch (psychotic)	0.98	0.12 (0.09 to 0.16)	<.001	Grond (ground)	0.98	0.10 (0.06 to 0.14)	<.001
15	Collega (colleague)	0.98	0.11 (0.07 to 0.15)	<.001	Aanvang (commencement)	0.98	0.11 (0.08 to 0.14)	.01
16	Spreekt (speaks)	0.97	0.12 (0.08 to 0.15)	<.001	Mede (also)	0.98	0.10 (0.07 to 0.14)	.001
17	Gehouden (obliged)	0.97	0.11 (0.07 to 0.15)	<.001	Dhr wilde (Mr wanted)	0.98	0.10 (0.06 to 0.14)	.001
18	Beoordelen (judge), verb	0.96	0.11 (0.07 to 0.15)	<.001	Liep (walked)	0.98	0.10 (0.06 to 0.14)	.006
19	Momenten (moments)	0.96	0.12 (0.08 to 0.15)	<.001	Geagiteerd (agitated)	0.96	0.10 (0.06 to 0.14)	.01
20	Somber (dejected)	0.95	−0.14 (−0.17 to −0.11)	<.001	cvd (not available)	0.96	0.10 (0.06 to 0.14)	.004

FASE 5/6:
MODEL EVALUATIE → VERTROUWEN DE EINDGEBRUIKERS DE VOORSPELLINGEN?



- Sample of Local Explanation predicting <u>high</u> risk of aggression



- Sample of Local Explanation predicing <u>low</u> risk of aggression

The *"Linear Model-Agnostic Explanations"* (LIME) method

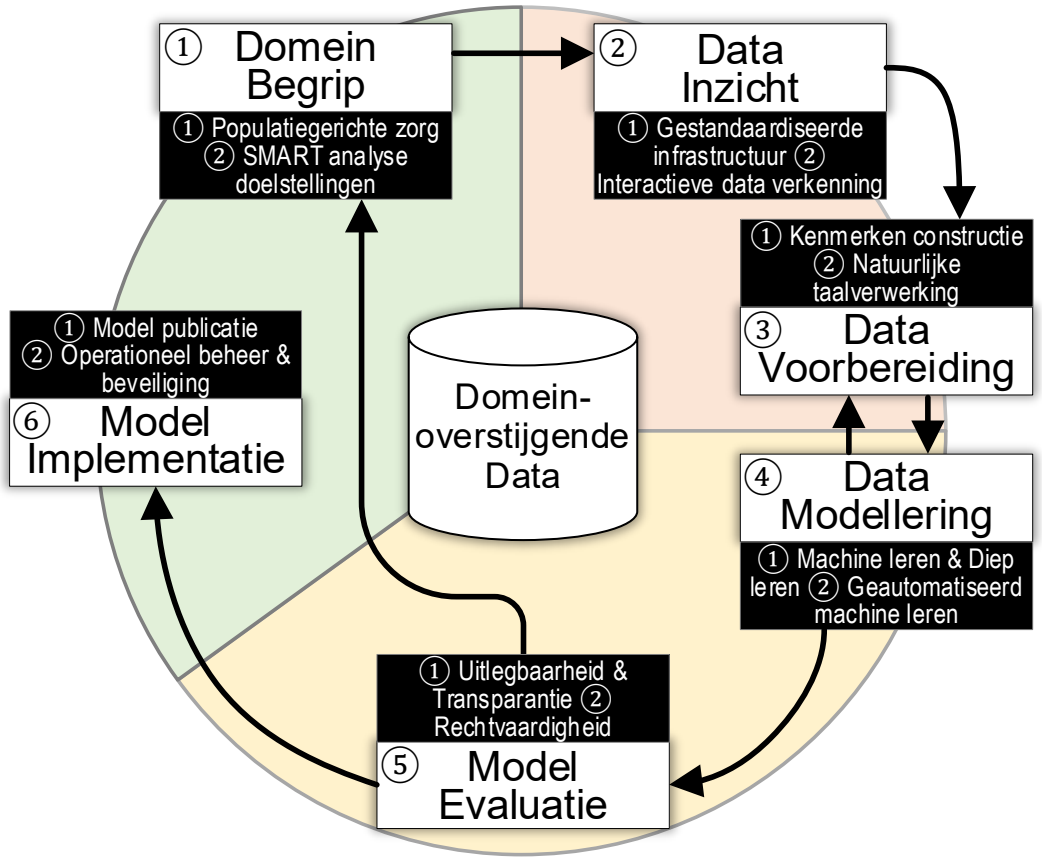
Fase 6/6:
Model Implementatie?

- Mosteiro,P., Rijcken,E., Zervanou,K., Kaymak,U., Scheepers,F., & Spruit,M. (2021). Machine Learning for Violence Risk Assessment Using Dutch Clinical Notes. *Journal of Artificial Intelligence for Medical Sciences, 2*(1–2), 44–54. [pdf] [online]
- Rijcken,E., Kaymak,U., Scheepers,F., Mosteiro,P., Zervanou,K., & Spruit,M. (2022). Topic Modeling for Interpretable Text Classification from EHRs. *Frontiers in Big Data, 5*, Section Data Mining and Management, 846930. [pdf] [online]
- …

Dank voor jullie aandacht 👍

Web: www.universiteitleiden.nl/medewerkers/marco-spruit

Contact: m.r.spruit@lumc.nl



Data Techniek Data Analyse e-Health Implementatie

TOEGIFT: EXTRACTIE VAN LEVENSSTIJL KENMERKEN

ROOKSTATUS, ALCOHOL- EN DRUGSGEBRUIK

Hielke Muizelaar, Peter van der Putten, & Marco Spruit (2023).
Exploring Dutch BERT Models for Extracting Lifestyle Characteristics from Medical Text. CLIN 2023.

EPD DATA SAMENVATTING

Type label	#Teksten	[1/4] Huidige gebruikers	[2/4] Voorm. gebruikers	[3/4] Niet-gebruikers	[4/4] Geen info
Roken	148.768	7.015 (4.72%)	32.230 (21.66%)	44.677 (30.03%)	64.846 (43.59%)
Drinken	143.166	16.017 (11.25%)	-	39.119 (27.32%)	87.940 (61.43%)
Drugs	147.999	1.443 (0.98%)	-	53.005 (35.81%)	93.551 (63.21%)

Voorbeeld tekst	Roken	Alcohol	Drugs
Patïent rookt, drinkt niet, gebruikt geen drugs	Huidige gebr.	Niet-gebruiker	Niet-gebruiker
Patient rookte vroeger, drinkt 1 bier/dag	Voormalige gebr.	Huidige gebr.	Geen info
Patient rookte vroeger, gebruikt nu dagelijks wiet	Voormalige gebr.	Geen info	Huidige gebr.

BERT-model Resultaten

Model	Macro F1-score Smoking	Macro F1-score Alcohol	Macro F1-score Drugs
String Matching	0.84	0.74	0.68
Standard Machine Learning (SGD)	0.85	0.71	0.60
HAGALBERT	0.66	0.54	0.43
RobBERT-HAGA	0.87	0.71	0.63
belabBERT-HAGA	0.48	0.64	0.57
MedRoBERTa.nl-HAGA	**0.93**	0.79	**0.77**
BioBERT (translated)	0.91	0.72	0.52
ClinicalBERT (translated)	0.92	**0.80**	0.61