



Universiteit
Leiden
The Netherlands

Natural language processing in healthcare: applications and value

Buchem, M.M. van

Citation

Buchem, M. M. van. (2024, December 11). *Natural language processing in healthcare: applications and value*. Retrieved from <https://hdl.handle.net/1887/4172376>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4172376>

Note: To cite this publication please use the final published version (if applicable).



Chapter 11

Samenvatting

Dit proefschrift onderzoekt de toepassing van natuurlijke taalverwerking (NLP) in de gezondheidszorg. Deel 1 beschrijft de ontwikkeling van verschillende NLP-modellen in diverse zorgsettings en bespreekt de bijbehorende uitdagingen. Deel 2 richt zich op het vaststellen van de toegevoegde waarde van NLP in de klinische praktijk en bespreekt twee pilotstudies.

Hoofdstuk 2 biedt een uitgebreide scoping review en een onderzoeksagenda voor de implementatie van *digital scribes* in de klinische praktijk. Digital scribes, die gebruik maken van automatische spraakherkenning (ASR) en NLP-technieken, zijn bedoeld om de werkdruk van klinici te verlichten door de klinische documentatie te automatiseren. Onze review benadrukt dat huidig onderzoek voornamelijk gericht is op technische validiteit, zonder voldoende aandacht te besteden aan klinische validiteit, bruikbaarheid of nut. Aanbevelingen voor toekomstig onderzoek omvatten het verbeteren van de ASR-nauwkeurigheid, het waarborgen van uitgebreide validatie en het verbeteren van de transparantie in rapportage om de kloof tussen onderzoek en praktische implementatie te overbruggen.

Hoofdstuk 3 evalueert het gebruik van NLP om het gebruik van acute zorg bij oncologiepatiënten die starten met chemotherapie te voorspellen, door gebruik te maken van klinische notities. Het vergelijkt deep learning-modellen met modellen die gebruik maken van handmatig geëxtraheerde taalfeatures en gestructureerde gezondheidsgegevens (SHD). De resultaten geven aan dat, hoewel SHD-modellen iets beter presteren dan NLP-modellen, beide benaderingen rendabel zijn. De studie benadrukt het potentieel van NLP in klinische toepassingen en wijst op risico bias over diverse patiëntengroepen, wat suggereert dat toekomstig onderzoek zich moet richten op het verbeteren van de generaliseerbaarheid van modellen en het aanpakken van deze bias.

Hoofdstuk 4 bespreekt de ontwikkeling en validatie van de Artificial Intelligence Patient-Reported Experience Measures (AI-PREM), een tool die is ontworpen om de analyse van vrije tekst patiëntervaringsgegevens te automatiseren met behulp van NLP (zie Figuur 1). De AI-PREM omvat een nieuw ontwikkelde vragenlijst met open vragen, een NLP-pijplijn voor het analyseren van reacties via sentimentanalyse en topic modeling, en een visualisatie-interface voor gemak-

kelijke interpretatie door zorgprofessionals. Het validatieproces bevestigde de mogelijkheid van de tool om relevante patiëntervaringsthema's en -sentimenten nauwkeurig te identificeren, wat de toepasbaarheid ervan suggereert ter ondersteuning van kwaliteitsverbetering in klinische settings.

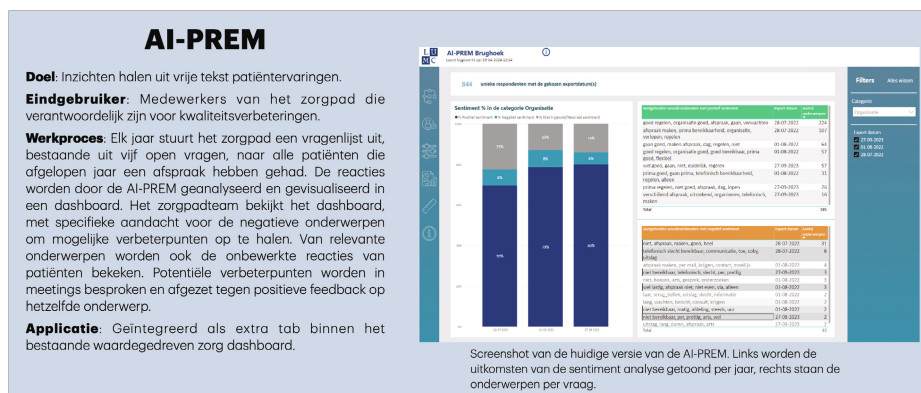
Hoofdstuk 5 presenteert de ontwikkeling van een model om mogelijke depressie bij kankerpatiënten te identificeren met behulp van NLP op patiëntberichten. De studie maakte gebruik van berichten van een beveiligd patiëntenportaal in een kankercentrum, waarbij gebruik werd gemaakt van logistische regressie, support vector machines en twee varianten van BERT-modellen (origineel en voorgetraind op Reddit) voor classificatie. De beste prestaties werden bereikt door de BERT-modellen. Dit geeft een behoorlijk goede effectiviteit aan in het detecteren van mogelijke depressie. De studie onderzocht ook de rechtvaardigheid en verklaarbaarheid van modellen, waarbij prestatieverschillen tussen demografische groepen aan het licht kwamen. Dit onderzoek onderstreept het potentieel van geavanceerde NLP-technieken voor vroege depressiedetectie in klinische settings, hoewel het ook de noodzaak benadrukt om bias en hun impact op verschillende patiëntengroepen zorgvuldig te overwegen.

Hoofdstuk 6 onderzoekt de publieke percepties van statines op Reddit, door met AI 10.233 posts en commentaren te analyseren. NLP-modellen en clusteringsalgoritmen categoriseerden de inhoud in 100 onderwerpen over zes thematische groepen. De studie suggereert dat AI waardevolle inzichten kan bieden in publieke opinies, waardoor zorgprofessionals misvattingen kunnen aanpakken en de naleving van statinegebruik kunnen verbeteren.

Hoofdstuk 7 onderzoekt de waarde van de AI-PREM-tool in de klinische praktijk, specifiek voor het brughoektumorzorgpad (zie Figuur 1). Door open en gesloten patiëntervaringsvragenlijsten te vergelijken, vond de studie dat de AI-PREM meer gedetailleerde patiëntfeedback bood en specifieke verbeterpunten identificeerde die gesloten vragenlijsten misten. De bevindingen benadrukken het potentieel van AI-PREM om patiëntgerichte zorg te verbeteren door middel van actiegericht feedback.

Hoofdstuk 8 onderzoekt de impact van het automatiseren van klinische documentatie met grote taalmodellen, met de nadruk op tijd en kwaliteit (zie Figuur 2). Geneeskundestudenten hebben 150 gesimuleerde gesprekken samengevat, hetzij handmatig, hetzij met een Nederlands digital scribe-systeem. Zij ontdekten dat volledig geautomatiseerde samenvattingen lagere kwaliteitsscores hadden vergeleken met handmatige samenvattingen, maar dat de kwaliteit verbeterde na het bewerken van geautomatiseerde samenvattingen. Bovendien kostte handmatig samenvatten meer tijd dan het bewerken van geautomatiseerde samenvattingen. De studie suggereert dat digital scribes de documentatietijd kunnen verminderen terwijl de kwaliteit van de verslagen hoog blijft, met verder onderzoek nodig in klinische omgevingen.

Concluderend: er is aanzienlijke vooruitgang in natural language processing (NLP) met vele potentiële toepassingen in de gezondheidszorg. Ondanks deze vooruitgang belemmeren uitdagingen zoals datakwaliteit, bias, klinisch relevante maten, en gebrek aan goede klinische studies de brede acceptatie. Er zijn gerichte inspanningen vereist, zoals meer klinische evaluaties, verbeterde rapportage en meer gestroomlijnde integratie van NLP in de klinische context. Dan kan het potentieel van NLP worden gerealiseerd voor het verbeteren van patiëntenuitkomsten, de ervaring van clinici en het verlagen van kosten.



Figuur 1: een beschrijving van de AI-PREM tool.



Figuur 2: een beschrijving van de Autoscriber tool.