

Herkennen van depressie met computationele taalanalyse

S. Koops, S.G. Brederoo, J.N. de Boer, A.E. Voppel, I.E.C. Sommer

- Achtergrond** Op dit moment missen we in de klinische praktijk een bruikbare biomarker voor het detecteren en differentiëren van depressie. Zo'n biomarker zou de spontane spraak kunnen zijn, waar we belangrijke informatie uit kunnen halen met geautomatiseerde taalanalyse.
- Doel** Een overzicht bieden van het zich snel ontwikkelende veld van geautomatiseerde taalanalyse bij depressie.
- Methode** Samenvatting van de huidige literatuur over kenmerken van depressie in gesproken taal.
- Resultaten** De huidige computationele modellen kunnen depressie detecteren met hoge nauwkeurigheid, wat deze goed toepasbaar maakt in diagnostische hulpmiddelen gebaseerd op automatische taalanalyse. Dergelijke hulpmiddelen zijn volop in ontwikkeling.
- Conclusie** Voor klinische implementatie staan nog enkele hindernissen in de weg. Zo wisselen de resultaten sterk per studie wegens grote variatie in gebruikte methodologie. Ook privacybescherming en ethische kwesties moeten we oplossen voordat we automatische taalanalyse kunnen implementeren.

Depressieve stoornissen (hierna genoemd: depressie) komen veel voor en gaan gepaard met ernstig lijden, een hoge kans op zelfdoding, hoge gezondheidszorgkosten en hoog werkverzuim.¹

Accurate en tijdige diagnose is van groot belang voor succesvolle behandeling, maar blijkt uitdagend. Veel gevallen worden bijvoorbeeld gemist.² Objectieve biomarkers voor het vaststellen van een diagnose ontbreken. De psychiater is voor de diagnostiek afhankelijk van door de patiënt gerapporteerde symptomen en zijn/haar eigen observatie; beide zijn per definitie subjectief. Een objectieve biomarker voor depressie zouden we kunnen vinden in automatische analyse van spraak. Onze spraak geeft veel informatie over hoe het met ons gaat en is een belangrijk onderdeel van zowel het psychiatrisch onderzoek als de anamnese. Spraak kunnen we gemakkelijk, goedkoop en niet-invasief verkrijgen en kunnen we ook longitudinaal frequent gebruiken met weinig belasting voor de patiënt.

Spraak van patiënten met een depressie wordt veelal als traag en eentonig ervaren. Zowel de akoestische kenmerken van spraak, d.w.z. de meer fysieke kenmerken, alsook de inhoud van de spraak zijn afwijkend bij depressie. Recente ontwikkelingen in natuurlijke taalverwerkingstechnieken (*natural language processing*; NLP) waarin patroonherkenning centraal staat, bieden

de mogelijkheid om zelfs de subtielste vormen van dergelijke spraakafwijkingen efficiënt en snel te detecteren. Met een narratieve review bieden wij een overzicht van het zich snel ontwikkelende veld van geautomatiseerde taalanalyse bij depressie.

Computationele taalanalyse bij depressie

Detectie met akoestische kenmerken

Veelal gebruikt men binnen computationele analyses akoestische taalkenmerken om depressie te detecteren. Dit zijn kenmerken die gemeten worden in het audiosignaal van spraak, zoals prosodische kenmerken (bijv. klemtoon en intonatie), timingkenmerken (bijv. spreektempo en pauzetijd), volume, stemkwaliteit, en andere frequentie- en spectrale kenmerken (bijv. toonhoogte). Een typische bevinding bij depressie is een afname in toonhoogte en toonhoogtebereik – passend bij monotone spraak.³ Kenmerken zoals stemkwaliteit en variatie in toonhoogte zijn niet alleen een indicator van de aanwezigheid van depressie, maar correleren ook met de ernst van depressie en (het uitblijven van) behandelingsrespons.^{1,3,4}

Wat betreft timinggerelateerde spraakkenmerken is er sprake van een lager spreektempo, verhoogde pauze- en reactietijd.^{1,5} Deze maten correleren ook significant met

AUTEURS

Sanne Koops, postdoctoraal onderzoeker, afd. Psychiatrie, UMC Groningen.

Sanne Brederoo, psycholoog en postdoctoraal onderzoeker, afd. Psychiatrie, UMC Groningen.

Janna de Boer, psychiater in opleiding, afd. Psychiatrie, UMC Utrecht.

Alban Voppel, postdoctoraal onderzoeker, afd. Psychiatrie, UMC Groningen.

Iris Sommer, psychiater en hoogleraar Cognitieve aspecten van neurologische en psychiatrische stoornissen, afd. Psychiatrie, UMC Groningen.

Correspondentie

Sanne Koops (s.schuite-koops@umcg.nl).

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 31-8-2022.

Citeren

Tijdschr Psychiatr. 2023;65(3):198-201

de ernst van depressie.⁶ Dit soort modellen halen hoge accuratesse tot zelfs 91% in het detecteren van depressie.⁷

Detectie met inhoudelijke taalkenmerken

Ook inhoudelijke aspecten van gesproken taal zijn bruikbaar in automatische taalanalyse om depressie te detecteren. Zo zijn er verschillen tussen individuen met en zonder depressie in onder andere syntactische (grammaticale) en lexicale (de woordkeuze betreffende) taaldimensies. Bij depressie is er bijvoorbeeld meer zelfreferentie, zoals het gebruik van de eerste persoon enkelvoud ('ik', 'me' en 'mij'),⁸ passend bij de zelfgerichte taaltheorie waarbij depressie is gekarakteriseerd door een (negatieve) zelffocus.⁹ Ook kenmerkend is veelvuldig gebruik van verleden-tijdsvervoegingen voor actiewerkwoorden; omgekeerde woordvolgorde voor onderwerp; minder gebruik van nadruk; korte en kleurloze uitspraken; tautologieën; herhalingen; en gebrek aan vergelijkingen.^{8,10,11} Taalkundige analyse is bij depressie veel op geschreven taal toegepast, bijvoorbeeld met sociale media.¹² Bij gesproken taal worden deze analyses uitgevoerd met getranscribeerde tekst van een spraakopname, wat inmiddels kan met geautomatiseerde methoden. Een veelbelovende NLP-techniek is het bestuderen van betekenis en coherentie in taal met 'semantic-space-modellen'.¹³ Deze modellen gaan uit van het idee dat de betekenis (semantiek) van woorden afhankelijk is van hun context. In coherente spraak zal een woord omringd worden door woorden die qua betekenis daaraan gerelateerd zijn.

Een andere methode om semantische informatie in taal te analyseren is de lexicongebaseerde benadering. Een lexicon bevat alle woorden die worden gebruikt binnen een specifiek onderwerp of een taal. Een depressielexicon bevat dus idealiter alle woorden die worden gebruikt door individuen met depressie. Onderzoek laat zien dat het combineren van semantic-spacemodellen zoals word2vec met lexicongebaseerde methoden nog accuratere depressieherkenning oplevert.¹⁴

Multimodale analyses

In multimodale analyses kan men spraakaspecten combineren met andere kenmerken zoals gezichtsuitdrukkingen of bewegingen, wat de prestatie van de modellen lijkt te verbeteren.^{15,16} De accuratesse van dergelijke modellen is wel afhankelijk van welke categorie vragen worden gesteld.¹⁶

Detectie op basis van smartphonedata

Met smartphones kan men spraakanalyse in een naturalistische setting toepassen, omdat men deze kan uitrusten met gezondheidsapplicaties om zowel fysieke als mentale gezondheid te verbeteren.

Men kan dan denken aan zelfmonitoring als terugvalpreventie middels een vorm van spraakbiofeedback,¹⁷ of de Engelstalige app *Psychologist in a Pocket* die via geautomatiseerde lexicongebaseerde tekstanalyse symptomen van depressie detecteert.¹⁸ Een voordeel van zulke gezondheidsapplicaties is dat ze laagdrempelig zijn en weinig schaamte of stigma oproepen bij het initiële hulpzoeken.

Een uitdaging is wel dat spraakopnames met een smartphone in verschillende omgevingen worden gemaakt, met slecht controleerbaar achtergrondgeluid, al lijken enkele specifieke spraakkenmerken hiertegen bestand.¹⁹ Ook passief verkregen spraakopnames zijn bruikbaar, waar vele mogelijkheden toe bestaan sinds de opkomst van stembedieningsassistenten (zoals Apples Siri, Amazons Alexa en Google assistant). Hier zijn echter wel uitdagingen op het gebied van de privacy. Desondanks is deze mobiele variant van spraakanalyse veelbelovend om een objectieve en gemakkelijke tool voor depressie-detectie te worden. Recent onderzoek door onze groep laat zien dat mensen met psychische klachten zo'n applicatie graag zouden gebruiken mits onduidelijkheden rondom privacy weggenomen worden.²⁰

Privacy en ethische overwegingen

Uiteraard komen er behoorlijke privacyproblemen bovendrijven bij apparaten die luisteren naar spraak en software gebruiken om die data te verwerken, bijvoorbeeld het risico op datalekken. Voor gebruikers is niet altijd duidelijk waar hun data voor worden gebruikt, door afwezigheid van eenvoudig leesbare documenten over privacybeleid.²¹ Zo was er in 2019 ophef rondom Apples Siri, toen bleek dat Apple-medewerkers luisterden naar passieve spraakopnames.

Machinelearningmodellen zouden informatie kunnen lekken over de individuele datasets waarop ze zijn getraind.²² Modellen zouden tevens op discriminerende wijze gebruikt kunnen worden of tot systematische meetfouten leiden. Een voorbeeld hiervan is software van Amazon, die een systematische meetfout ten nadele van vrouwen bleek te hebben.²³ Daarnaast moet men ethische dilemma's onderzoeken bij de implementatie van modellen of applicaties die gebruikmaken van automatische taalanalyse, met name als dit data betreft die van sociale media of smartphones worden gehaald. Denk bijvoorbeeld aan potentieel discriminerend gebruik van de modellen door zorgverzekeraars of werkgevers.³

Ook bij spraakopnames tijdens bijvoorbeeld intakes zijn dit problemen die men nauwlettend in de gaten moet houden en tegengaan, zij het mogelijk in mindere mate doordat er geen sprake is van passieve opnameapparatuur of dataverwerking in de 'cloud'. Een belangrijke eerste stap is om enkel zogenaamde 'meta-data' van spraak op te slaan, dus de van spraak afgeleide maten, maar niet de verbatim spraak zelf.

Gouden standaard

Een belangrijke beperking is de gouden standaard waarmee de prestatie van taalanalysemodellen wordt vergeleken: het subjectieve diagnostische proces dat we door de taalanalyse zouden moeten verbeteren. Er is echter geen andere standaard beschikbaar om de resultaten mee te vergelijken. Taalanalyse zou een mooi startpunt kunnen zijn om psychiatrische diagnoses te verfijnen: initieel ontwikkeld om op hetzelfde niveau te komen als het huidige diagnostische proces om het daarna te overstijgen. Zo kon onze onderzoeksgroep binnen een groep patiënten met een schizofreniespectrumstoornis subtypen onderscheiden met specifieke ziekte-trajecten, symptomen en cognitief functioneren.²⁴ Een goed ontwikkelde spraakbiomarker kan ook de categorisatie van patiënten met depressie verfijnen en relevante informatie bieden over hun symptomen, functioneren, prognose en dus mogelijke behandelopties.

DISCUSSIE

Automatische taalanalyse lijkt een betrouwbare en accurate methode om depressie te detecteren. Machinelearningmodellen halen hoge accuratesse in het detecteren

van depressie. Spraak biedt een objectieve en gemakkelijke biomarker in een werkveld waar objectieve biomarkers tot op heden niet voorhanden waren. Door bijvoorbeeld spraakopnames tijdens klinische interviews te gebruiken voor taalanalyse, kan de methode een goede aanvulling zijn op de reguliere gezondheidszorg. Spraak kan ook toegankelijk en discreet worden opgenomen en geanalyseerd middels smartphoneapplicaties.

Wel zorgt de grote variatie aan gebruikte technieken en rekenmodellen voor wisselende resultaten. We geven in dit artikel geen uitgebreide uitleg van de vele technieken, maar willen wel noemen dat deze grote invloed hebben op de resultaten. Belangrijk is het behalen van een goede generaliseerbaarheid van de modellen. Het meeste onderzoek tot nu toe is gericht op de analyse van akoestische kenmerken van spraak. Een combinatie van akoestische en inhoudelijke kenmerken, of een combinatie met niet-spraakgerelateerde data zoals gezichtsuitdrukkingen, lijkt de accuratesse van de bestaande modellen verder te verbeteren.

Er zijn nog enkele obstakels voordat we deze techniek in de dagelijkse praktijk kunnen gebruiken. Allereerst moet alle dataverzameling en data-analyse voldoen aan strenge privacy-eisen om te voorkomen dat individuele data ongevraagd gebruikt worden. Ook missen er longitudinale studies, waardoor belangrijke informatie ontbreekt over de mogelijke invloed van medicatie of andere behandeling op spraak. Veel patiënten met depressie gebruiken immers psychotrope medicatie, die invloed uitoefent op spraak.^{4,25}

In veel studies probeert men een patiënt met depressie te onderscheiden van iemand zonder depressie, of te differentiëren tussen categorieën van depressie. Het zou echter veel klinische waarde hebben als spraakmarkers kunnen differentiëren tussen verschillende psychiatrische ziektebeelden, zoals tussen depressie en angststoornissen of psychose, en de vele andere aandoeningen die spraak kunnen beïnvloeden zoals het chronischevermoeidheidssyndroom of verhoogde stress door heftige gebeurtenissen. Ook zou de spraakmarker potentieel kunnen differentiëren tussen depressie op oudere leeftijd en aandoeningen zoals de ziekte van Parkinson, de ziekte van Alzheimer of frontotemporale dementie.

CONCLUSIE

Spraakkenmerken worden steeds bruikbaar als effectief diagnostisch hulpmiddel voor depressie. Dit kan leiden tot snellere en objectievere diagnose, en daarmee een snellere start van doelgerichte en gepersonaliseerde behandeling. Ook tijdige terugvalpreventie en tijdige interventie behoren tot de waardevolle mogelijkheden.

LITERATUUR

- 1 Mundt JC, Vogel AP, Feltner DE, e.a. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry* 2012; 72: 580-7.
- 2 Faisal-Cury A, Ziebold C, Rodrigues DMO, e.a. Depression under-diagnosis: Prevalence and associated factors. A population-based study. *J Psychiatr Res* 2022; 151: 157-65.
- 3 Low DM, Bentley KH, Ghosh SS. Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investig Otolaryngol* 2020; 5: 96-116.
- 4 Silva WJ, Lopes L, Galdino MKC, e.a. Voice acoustic parameters as predictors of depression. *J Voice* 2021.
- 5 Liu Z, Kang H, Feng L, e.a. Speech pause time: A potential biomarker for depression detection. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; 2017. p. 2020-5.
- 6 Yamamoto M, Takamiya A, Sawada K, e.a. Using speech recognition technology to investigate the association between timing-related speech features and depression severity. *PLoS One* 2020; 15: e0238726.
- 7 Espinola CW, Gomes JC, Pereira JMS, e.a. Vocal acoustic analysis and machine learning for the identification of schizophrenia. *Research on Biomedical Engineering* 2021; 37: 33-46.
- 8 Tackman AM, Sbarra DA, Carey AL, e.a. Depression, negative emotionality, and self-referential language: A multi-lab, multi-measure, and multi-language-task research synthesis. *J Pers Soc Psychol* 2019; 116: 817.
- 9 Jarrold W, Javitz HS, Krasnow R, e.a. Depression and self-focused language in structured interviews with older men. *Psychol Rep* 2011; 109: 686-700.
- 10 Trifu RN, Nemeş B, Bodea-Haşegan C, e.a. Linguistic indicators of language in major depressive disorder (MDD). An evidence based research. *J Evid Based Psychother* 2017; 17.
- 11 Xu S, Yang Z, Chakraborty D, e.a. Automated verbal and non-verbal speech analysis of interviews of individuals with schizophrenia and depression. *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; 2019. p. 225-8.
- 12 Guntuku SC, Yaden DB, Kern ML, e.a. Detecting depression and mental illness on social media: An integrative review. *Curr Opin Behav Sci* 2017; 18: 43-9.
- 13 de Boer JN, Voppel AE, Begemann MJH, e.a. Clinical use of semantic space models in psychiatry and neurology: A systematic review and meta-analysis. *Neurosci Biobehav Rev* 2018; 93: 85-92.
- 14 Sugathadasa K, Ayesha B, de Silva N, e.a. Synergistic union of word2vec and lexicon for domain specific semantic similarity. *IEEE International Conference on Industrial and Information Systems (ICIIS)*; 2017. p. 1-6.
- 15 Williamson JR, Godoy E, Cha M, e.a. Detecting depression using vocal, facial and semantic communication cues. *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*; 2016. p. 11-8.
- 16 Guohou S, Lina Z, Dongsong Z. What reveals about depression level? The role of multimodal features at the level of interview questions. *Information & Management* 2020; 57: 103349.
- 17 Braun S, Annovazzi C, Botella C, e.a. Assessing chronic stress, coping skills, and mood disorders through speech analysis: A self-assessment 'voice app' for laptops, tablets, and smartphones. *Psychopathology* 2016; 49: 406-19.
- 18 Cheng PGF, Ramos RM, Bitsch Jü, e.a. Psychologist in a pocket: Lexicon development and content validation of a mobile-based app for depression screening. *JMIR Mhealth Uhealth* 2016; 4: e88.
- 19 Huang Z, Epps J, Joachim D. Speech landmark bigrams for depression detection from naturalistic smartphone speech. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; 2019. p. 5856-60.
- 20 Brederoo SG, Nadema FG, Goedhart FG, e.a. Implementation of automatic speech analysis for early detection of psychiatric symptoms: What do patients want? *J Psychiatr Res* 2021; 142: 299-301.
- 21 Sunyaev A, Dehling T, Taylor PL, e.a. Availability and quality of mobile health app privacy policies. *J Am Med Inform Assoc* 2015; 22: e28-33.
- 22 Shokri R, Stronati M, Song C, e.a. Membership inference attacks against machine learning models. *IEEE Symposium on Security and Privacy (SP)*; 2017. p. 3-18.
- 23 Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters* October 11, 2018. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.
- 24 Oomen PP, de Boer JN, Brederoo SG, e.a. Characterizing speech heterogeneity in schizophrenia-spectrum disorders. *J Psychopathol Clin Sci* 2022; 131: 172-81.
- 25 de Boer JN, Voppel AE, Brederoo SG, e.a. Language disturbances in schizophrenia: The relation with antipsychotic medication. *NPJ Schizophrenia* 2020; 6: 24.