

CHATBOT



Leerkrachtenhandleiding

Auteur: Natacha Gesquière

Coauteurs: Veronique Hoste, Cynthia Van Hee, Francis wyffels



“AI IS THE NEW ELECTRICITY”.

–ANDREW NG

Versie 1.0

Natacha Gesquière

Chatbot

Versie 1.0

Coauteurs:

Veronique Hoste, Cynthia Van Hee, Francis wyffels

CC BY-SA, 2021, Natacha Gesquière

Coauteurs: Veronique Hoste, Cynthia Van Hee, Francis wyffels

Corrector: Annick Dehennin

Illustrator: Margot De Saegher

Chatbot is een project van Dwengo vzw, UGent IDLab-AIRO en UGent LT³.

Chatbot maakt deel uit van AI Op School.

AI Op School werd in 2020 bekroond met een M-Award voor Mediawijs Educatief Project.

Het project KIKS van AI Op School won in 2020 de Koningin Paolaprijs voor het Onderwijs.

Deze handleiding kwam tot stand met (financiële of logistieke) steun van de Universiteit Gent, Dwengo vzw, de Provincie Oost-Vlaanderen, en de Vlaamse Overheid.

D/2021/Dwengo/1

ISBN 9789464441000

NUR 950

Eerste druk, augustus 2021



INHOUDSOPGAVE

1	Artificiële intelligentie	13
2	Geschiedenis van AI	25
3	Natural Language Processing	35
4	Chatbot	51
5	Sentimentanalyse	63
6	Cyberpesten	81
7	Persoonskenmerken	87
8	Gesproken taal	107
9	Ethiek	113

10 De notebooks	125
11 Eindtermen	129
Index	155

Versie 1.0

Met dank aan

Jeroen De Gussem (UGent),

Tom Neutens (UGent),

Marianne Petit (Sint-Bavohumaniora),

Zimcke Van de Staey (Dwengo vzw).

Versie 1.0

INLEIDING

Waar taal en technologie samenkomen, ontstaat het domein van *Natural Language Processing* (NLP). Met technieken uit *machine learning* kunnen we computers op een intelligente manier laten werken met taal; denk bijvoorbeeld aan het trainen van modellen om tekst, afbeeldingen of geluidsfrequenties te herkennen. Bekende NLP-toepassingen van *machine learning* zijn bijvoorbeeld vertaalprogramma's, spraakbesturingssystemen of virtual assistants zoals Siri, en chatrobots.

We buigen ons over vragen als “Hoe kunnen we automatisch teksten begrijpen, vertalen en zelfs schrijven?”, “Kan een computer emoties of ironie herkennen in een tekst?”, en vooral “Hoe belangrijk zijn data?”. Leerlingen van de tweede en de derde graad secundair onderwijs krijgen inzicht in het domein NLP en zetten hun eerste stappen in de wereld van het programmeren met Python.

In het domein van *Natural Language Processing* doen onderzoekers een beroep op algoritmen om teksten onder de loep te nemen. Ze gebruiken daarvoor zowel kennisgebaseerde als datagebaseerde kunstmatige intelligentie. Zo kunnen ze via sentimentanalyse nagaan of mensen zich positief of negatief uitlaten over bepaalde producten, bedrijven of personen op sociale media. Aan de hand van de tekststijl proberen ze te

achterhalen wie een bepaalde tekst schreef. Ze onderzoeken ook hoe ze persoonlijkheidstypes kunnen onderscheiden op basis van taal. Een goede kennis van taalkunde, computerwetenschappen en psychologie zijn daarbij onontbeerlijk; NLP is een interdisciplinair domein.

De ontwikkelingen in de taaltechnologie leveren nieuwe tools om fake nieuws op te sporen, criminaliteit te bestrijden en de gezondheidszorg efficiënter te maken; ook in de bedrijfswereld zijn er tal van toepassingen.

Dit verhaal roept ook ethische vragen op. Voor het onderzoek binnen taaltechnologie wordt gebruik gemaakt van data, bijvoorbeeld boodschappen op sociale media en chatberichten. Maar hoe zit het dan met privacy? Vinden we het toelaatbaar dat iemands persoonlijkheid wordt nagegaan op basis van de teksten die hij of zij schrijft? Vinden we het wenselijk hoe deze kennis vervolgens wordt aangewend? Ook de onderzoekers worstelen met deze vragen. Sommigen ontwikkelen daarom bijvoorbeeld algoritmes die teksten depersonaliseren.

Met deze handleiding voor de leerkracht willen we de leerkrachten die met Chatbot aan de slag gaan in de klas, achtergrondkennis geven over de leerinhouden van dit project.

Achteraan vind je de eindtermen die aan bod komen.

ARTIFICIËLE INTELLIGENTIE

Kunstmatige (KI) of artificiële intelligentie (AI) komt volop aan bod in de media. Populaire thema's zijn: chatbots, zelfrijdende auto's, virtuele assistenten, privacy en het verlies aan jobs.

AI dringt meer en meer onze levens binnen en leidt nu al tot hervormingen in de dienstensector en de industrie.

Stanford professor Andrew Ng noemt AI zelfs de nieuwe elektriciteit.

Discussie

Wat wordt er bedoeld met de nieuwe elektriciteit?

1.1 Intelligentie

Meer dan 2000 jaar geleden waren Socrates en Plato al geboeid door de menselijke intelligentie, maar tot op heden is er nog steeds geen algemeen gangbare definitie van intelligentie.

Enkele voorbeelden van hoe intelligentie gedefinieerd wordt (Legg & Hutter, 2007; Tegmark, 2017):

- “Het vermogen om te leren, om te kunnen inschatten en om te kunnen fantaseren” (16de eeuw, dokter en

psycholoog Huarte).

- “Intelligentie is de globale geschiktheid van een individu om doelbewust te handelen, om rationeel te denken en om doeltreffend om te gaan met de omgeving” (1958, psycholoog Wechsler, de grondlegger van de alom gebruikte IQ-testen WAIS en WISC).
- “Intelligentie is een heel algemene mentale geschiktheid die o.a. het vermogen inhoudt om te redeneren, te plannen, problemen op te lossen, abstract te denken, complexe ideeën te begrijpen, snel te leren en te leren uit ervaring” (1997, 52 experten).
- “Verstandelijk vermogen” (Van Dale).
- “Het vermogen om kennis en vaardigheden te verwerven en toe te passen” (Oxford English Dictionary).
- “Intelligentie meet het vermogen van een *agent* om doelen te bereiken in een breed spectrum van omgevingen” (2006, computerwetenschappers Hutter en Legg). Zo'n *agent* kan bv. een mens of een dier zijn, of een AI-systeem.
- “Intelligentie is het vermogen om complexe doelen te bereiken” (Tegmark, 2017).

De laatste twee definities omvatten, juist door hun breedheid, veel eigenschappen: het vermogen om kennis te verwerven, om zich aan te passen, om te begrijpen ...

1.2 AI

AI heeft al succesvolle toepassingen: gezichtsherkenning in China om toegang te verlenen tot gebouwen, AI-systemen die specialisten evenaren in de diagnose van huidkanker, AlphaGo Zero dat zelf go leerde spelen enkel door te spelen tegen zichzelf (Steels et al., 2017).

AlphaGo Zero is het resultaat van *reinforcement learning*, wat verder aan bod komt.

Een definitie geven van kunstmatige intelligentie, iets dat in zoveel verschillende domeinen wordt ingezet, is geen sinecure. Dus behalve van intelligentie, zijn er ook van AI al tal van definities gegeven:

- “Kunstmatige intelligentie verwijst naar systemen die intelligent gedrag vertonen door hun omgeving te analyseren en – in zekere mate zelfstandig – actie te ondernemen om specifieke doelstellingen te verwezenlijken. Op KI gebaseerde systemen kunnen louter softwarematig zijn en actief zijn in de virtuele wereld (bijvoorbeeld stemgestuurde assistenten, software voor beeldanalyse, zoekmachines en systemen voor spraak- en gezichtsherkenning), maar KI kan ook in hardware-apparaten worden geïntegreerd (bijvoorbeeld geavanceerde robots, zelfrijdende auto’s, drones of toepassingen van het internet der dingen)”, dit is de definitie die de Europese Commissie hanteerde in 2018 (PwC, 2018).
- “Artificiële intelligentie is niet-biologische intelligentie” (Tegmark, 2017).
- “AI verwijst naar computers die de dingen doen die men met verstand doet. Sommige van deze zaken (bv. redeneren) worden doorgaans omschreven als ‘intelligent’. Andere (bv. zicht) niet. Maar allemaal vereisen ze psychologische vaardigheden die het voor een mens en een dier mogelijk maken om hun doelen te bereiken, zoals perceptie, associatie, voorstelling, planning en het controleren van een motor” (Boden, 2016).

Discussie

Bedenk zelf een definitie van kunstmatige intelligentie.

Hoewel er ongetwijfeld spannende ontwikkelingen zijn

op het vlak van AI, staat de kunstmatige intelligentie vooralsnog mijlenver af van menselijke intelligentie.

Een mens kan behoorlijk veel bijleren uit een beperkt aantal voorbeelden, omdat hij ook zijn reeds opgedane kennis en vaardigheden benut. Een mens is bovendien zeer flexibel: we zijn in staat ons snel aan te passen aan veranderende omstandigheden en adequaat te reageren op onverwachte gebeurtenissen. Vandaag de dag kan een computer dat niet. Een AI-systeem is gericht op het invullen van een bepaalde taak en is maar goed in datgene waarvoor het ontworpen is.

Brede AI, enge AI en AGI

Bij enge AI (*narrow AI*) gaat het over AI-systemen die de specifieke taken doen waarvoor ze ontworpen werden, bv. een spamfilter.

Brede AI (*general AI*) betreft systemen die al meer diverse taken aankunnen, denk bv. aan wat men zich kan voorstellen bij een robotbutler.

Bij de zogenaamde *Artificial General Intelligence (AGI)* gaat het over computersystemen met dezelfde capaciteiten als een mens. We hebben vooralsnog niet de kennis om *AGI*-systemen te ontwerpen. Voorlopig zit AI nog maar op het niveau van de enge AI.

Discussie

Vergelijk menselijke intelligentie met artificiële intelligentie.

1.3 Soorten AI

1.3.1 Kennis- en datagebaseerde systemen

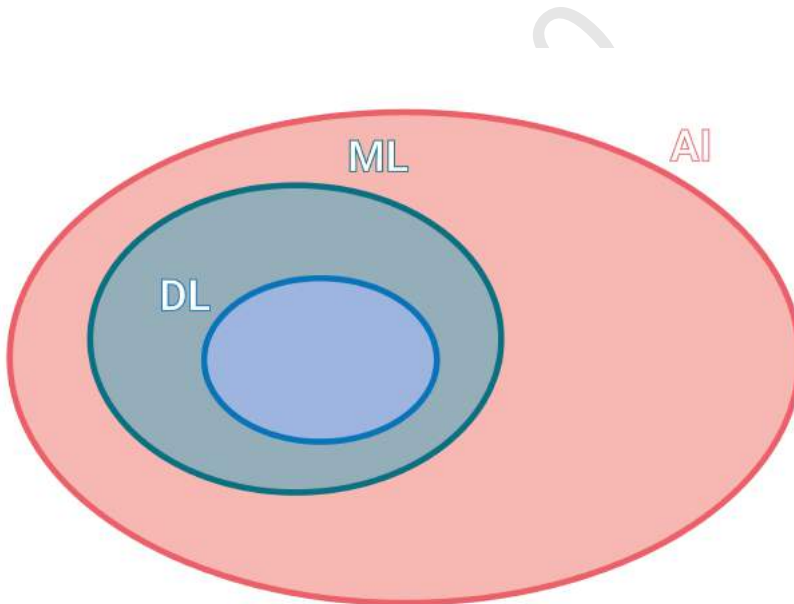
Om AI in te zetten voor een bepaald probleem, wordt informatie over dat probleem aan de computer gegeven. Vervolgens verwerkt het AI-systeem deze informatie en komt er een output.

Eerst moet zo'n AI-systeem ontwikkeld worden door het programmeren van algoritmes waarin expertenken-

nis vervat is of a.d.h.v. lerende algoritmes (Steels et al., 2017):

- Kennisgebaseerde systemen, die met regels worden geprogrammeerd. Dit houdt in dat men de kennis van menselijke experts zoveel mogelijk in regels probeert te gieten.
- Databasebaseerde systemen, die met lerende algoritmes worden geprogrammeerd en ook *machine learning*-systemen worden genoemd. Met statistische methodes worden patronen in relevante data opgespoord en dan gebruikt om nieuwe problemen op te lossen.

De chatbot ELIZA en lexicongebaseerde sentimentanalyse zijn voorbeelden van kennisgebaseerde systemen (zie verder in de hoofdstukken Chatbot en Sentimentanalyse). Gezichtsherkenning op Facebook en Google Translate zijn databasebaseerde systemen.



Figuur 1.1: Soorten AI.

Bij kennisgebaseerde systemen bestaat het algoritme uit expliciet geprogrammeerde regels. Men botst daarbij al snel op de grenzen van deze systemen. Voor een autonome wagen zou dat bijvoorbeeld betekenen dat men alle mogelijke scenario's zou moeten kennen.

1.3.2 Machine learning en deep learning

Machine learning (ML) is een populair en succesvol onderdeel van de artificiële intelligentie. Bij *machine learning* gaat men vooral proefondervindelijk te werk, maar op een wiskundig onderbouwde manier, en gebruikmakend van principes uit de wiskundige statistiek (Chollet, 2018).

Een *machine learning*-systeem verwerft met lerende algoritmes kennis uit data met de bedoeling uitkomsten te kunnen voorspellen betreffende nieuwe data. Deze voorspellingen worden gedaan met een bepaalde zekerheid. Zulke ML-modellen zullen dus nooit honderd procent accuraat zijn.

De nieuwe data moeten wel gelijkwaardig zijn aan de aangeboden data.

Lerende algoritmes

Lerende algoritmes zijn algoritmes waarbij het ML-systeem zelf gaandeweg aanpassingen doet aan de aanwezige parameters tijdens het leerproces, om geleidelijk aan te komen tot betere prestaties.

Voorspellen betekent bijvoorbeeld dat er uit voorbije tendensen cijfers voor de toekomst gegenereerd worden of dat een object bij een bepaalde categorie (klasse genoemd) wordt ingedeeld. Een ML-systeem neemt zijn beslissingen dus niet op basis van vooraf in detail geprogrammeerde instructies.

Binnen het machinaal leren bekleedt *deep learning* (DL) een prominente plaats (zie Figuur 1.1). DL heeft een revolutie veroorzaakt in het veld van ML door opmerkelijke resultaten te boeken op het vlak van spraak- en beeldherkenning, wat vaardigheden zijn waarvan de computer de menselijke prestaties tot op vandaag moeilijk kan evenaren. DL speelt behalve in de spraakherkenning ook een grote rol in andere domeinen van taaltechnologie.

In plaats van categorieën spreekt men in deze context eerder van klassen.

Het voorspellen gebaseerd op tendensen is een regressieprobleem. Het voorspellen van een klasse is een classificatieprobleem.

Voorbeelden: de prijs van een appartement voorspellen of beslissen of een e-mail al dan niet spam is.

1.3.3 Generaliseren bij ML

Een belangrijk doel van ML is om modellen te bekomen die goed kunnen generaliseren, m.a.w. modellen die goed presteren op ongeziene data. De generalisatie vormt steeds een uitdaging.

Stel dat men een spamfilter wilt trainen. Een spamfilter moet e-mails die spam zijn leren onderscheiden van e-mails die geen spam zijn.

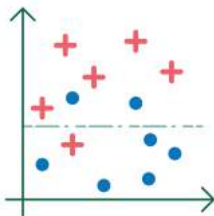
Zo'n ML-systeem wordt getraind aan de hand van een zogenaamde gelabelde trainingset: een groot aantal e-mails waarbij vermeld wordt of ze spam zijn of niet.

Na de training heeft men een model ter beschikking om nieuwe e-mails te beoordelen als zijnde spam of geen spam (zie Figuur 1.5). Dit model zal niet perfect zijn. Er zullen nog steeds spam e-mails door de mazen van het net glippen en sommige e-mails zullen onjuist in het spam postvak terechtkomen.

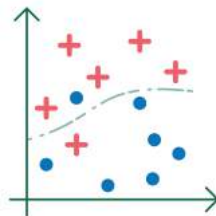
Hoe verloopt zo'n training?

Bij het begin van de training kiest het algoritme willekeurige waarden voor de parameters van het systeem. Het is dus niet verwonderlijk dat het netwerk initieel niet goed presteert. Gedurende de training past het algoritme de waarden van de parameters aan, gebaseerd op de gelabelde voorbeelden, en verbeteren de resultaten.

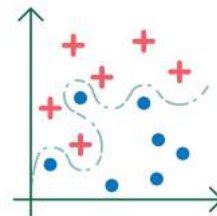
In de figuren stellen de kruisjes spam voor en de bolletjes e-mails die geen spam zijn.



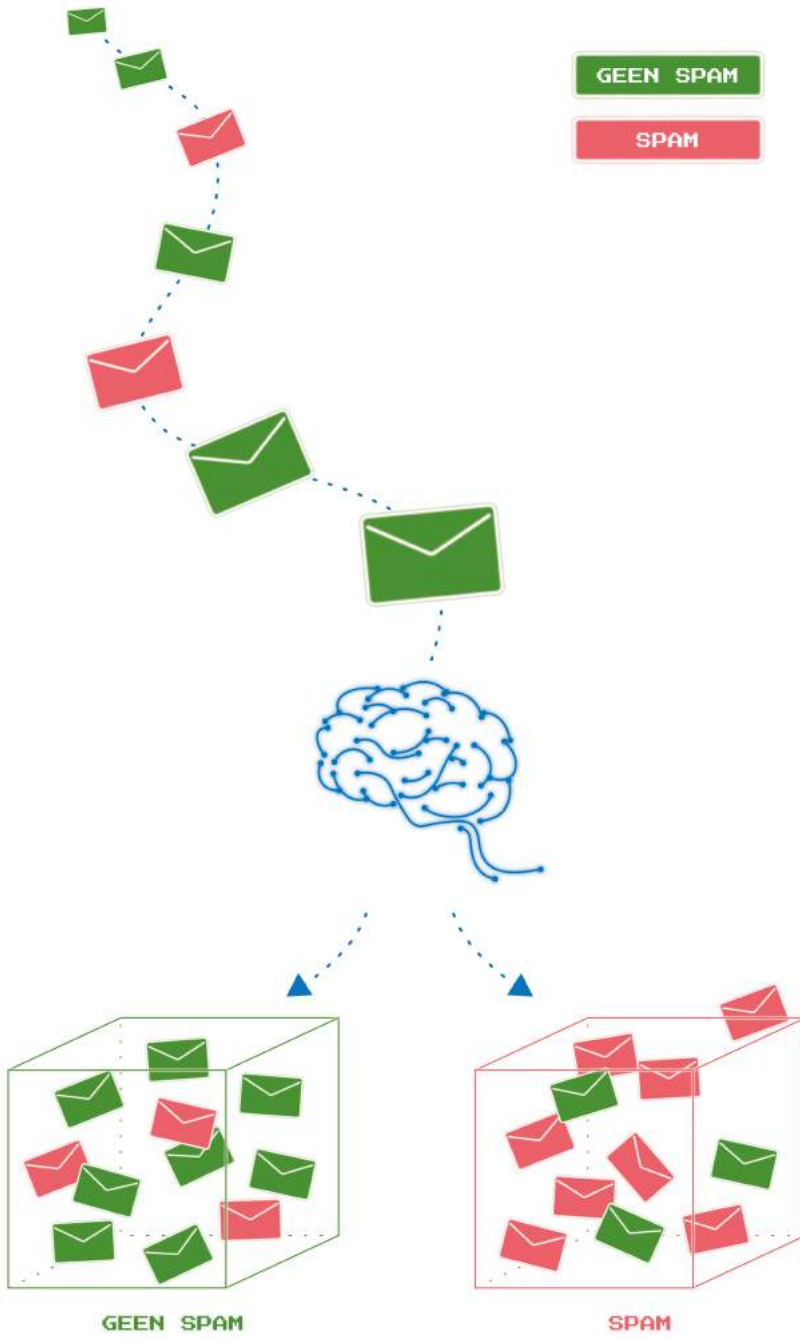
Figuur 1.2: Underfitting.



Figuur 1.3: Optimaal.



Figuur 1.4: Overfitting.



Figuur 1.5: Spamfilter.

Bij het trainen heeft het ML-systeem de twee categorieën van e-mails uit de trainingset van elkaar gescheiden; dit wordt in de figuren 1.2, 1.3, 1.4 gevisualiseerd door de scheidingslijn. De e-mails boven de scheiding worden geclassificeerd als spam en de e-mails eronder als geen spam.

In Figuur 1.2 doet het ML-systeem het nog niet zo goed. Het heeft de patronen om spam te onderscheiden van niet-spam nog niet herkend. Men gaat nog verder met trainen. Het systeem leert voortdurend meer.

In Figuur 1.3 is de scheiding beter. Daar zijn slechts twee e-mails van de trainingset bij de verkeerde soort ingedeeld. Men kan toch nog verder trainen, maar men moet opletten voor *overfitting*.

Op een bepaald moment in het trainingsproces wordt de optimale scheiding bereikt. Als men dan nog verder traint, dan zal het ML-systeem irrelevante kenmerken in de e-mails beginnen mee te nemen in zijn beoordeling. Het zal daardoor nog beter presteren op de e-mails van de trainingset, zoals in Figuur 1.4, maar niet beter op nieuwe e-mails. In dat geval is de situatie van Figuur 1.3 de optimale en zegt men dat het systeem in de situatie van Figuur 1.4 *overfit*.

Het komt er dus op aan om de architectuur van het systeem zo eenvoudig mogelijk te kiezen en op het juiste moment te stoppen met trainen om een zo goed mogelijk presterend model te bekomen. Er bestaan ook nog andere regularisatietechnieken om *overfitting* tegen te gaan, zoals data-augmentatie.

1.3.4 *Supervised, unsupervised en reinforcement learning*

Binnen ML onderscheidt men verschillende types van leren:

- *Supervised learning*: Het systeem leert uit een dataset waarbij elk gegeven bestaat uit twee componenten: een input gekoppeld aan een label (de ver-

Men zegt dat het systeem *underfits* in Figuur 1.2 omdat een beter presterend model bekomen kan worden. Het systeem in Figuur 1.2 heeft nog niet alle relevante patronen in de trainingdata ontdekt.

Op een bepaald moment in de training merkt men op dat de generalisatie niet meer verbetert. Er wordt enkel nog vooruitgang geboekt bij de trainingset. Het model *overfits*. Een te complex systeem of een systeem met teveel parameters zal sneller overfitten.

wachte output). Het labelen van de voorbeelden gebeurt vaak manueel door mensen, men noemt dat annoteren. Het systeem voert een algoritme uit dat er geleidelijk aan voor zorgt dat het systeem focust op relevante patronen in de data.

- *Unsupervised learning*: Hier bevat de dataset geen labels. Het AI-systeem moet op zoek naar kenmerken bij de verschillende voorbeelden en moet ze zo, door het ontdekken van patronen, verdelen in klassen. Men kan het systeem bv. ongelabelde foto's aanbieden van appels en peren. Het systeem gaat op zoek naar patronen om zo het onderscheid tussen de twee soorten fruit te kunnen maken.
- *Reinforcement learning*: Hier streeft het AI-systeem naar een beloning. Om bv. goed te worden in een bepaald videogame, gaat het AI-systeem het spel heel veel spelen en daaruit leren welke acties het beter vermijdt en welke het best onderneemt om te kunnen winnen. Zo werd Google DeepMind's AlphaGo Zero in 2017 via reinforcement learning een topspeler in go, nog beter dan AlphaGo die de beste menselijke speler eerder al versloeg.

Als men een opdracht uitvoert in het kader van reCAPTCHA, is men vaak aan het annoteren (bv. huisnummers, verkeerslichten en moeilijk leesbare woorden). M.b.v. reCAPTCHA werd het archief van The New York Times, met artikels sinds 1851, gedigitaliseerd.

Labelen van data

Analyseer je eigen Twitterberichten en leer aan computers om emoties te begrijpen.

<https://www.iedereenwetenschapper.be/projects/vertel-het-met-emoji039s>

Speel games en je helpt mee om vertaalsoftware te verbeteren.

<https://www.iedereenwetenschapper.be/projects/leer-computers-vertalen>

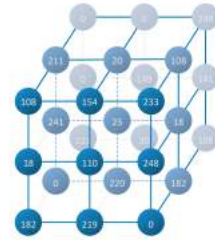
1.4 Voorstellen van de data

Om met AI een probleem te kunnen oplossen, geeft men via aangeboden data informatie over een bepaald probleem aan de computer. Deze data moeten echter in een voor het probleem geschikte vorm aangeboden worden. Men heeft m.a.w. een representatie van het probleem nodig waarmee de computer aan de slag kan. Vervolgens verwerkt het AI-systeem de informatie en komt er een output.

De data worden dus voorbereid voordat men ze aan het AI-systeem geeft. Een AI-systeem kan immers niets aanvangen met een foto, een tekst of een geluidsfragment als die niet op de juiste manier gerepresenteerd worden. Concreet werkt men met tensoren, dit zijn meerdimensionale rasteren die getallen bevatten, of anders gezegd matrices of ‘samenstellingen’ ervan (zie Figuur 1.6).

Voor een grijswaardenafbeelding volstaat één matrix, voor een kleurenfoto (zie Figuur 1.7) zijn er drie matrices nodig - een voor de rode, een voor de groene en een voor de blauwe tinten -, voor een film zijn er vier nodig. Voor een tekst gebruikt men vectoren, a.h.w. een lijst van getallen of anders gezegd één matrix met slechts een rij of kolom.

Dus ook voor het werken met teksten worden representaties met getallen gebruikt. Men kan dat bv. als volgt doen: alle woorden die in de data voorkomen, worden opgelijst. Deze woordenlijst wordt gegeven aan de computer. Elk woord in de woordenlijst kan dan worden voorgesteld door een rij van nullen en een één; bv. het eerste woord in de lijst door 1000... en het vierde woord door 000100... Om aan de computer kenbaar te maken welke van de woorden in de woordenlijst voorkomen in een bepaalde klantenreview, kan dan bv. een rij nullen en enen, zoals 0010111010000001000000100...0,



Figuur 1.6: Raster.



Figuur 1.7: Tensor van een kleurenafbeelding.

Er zijn nog andere manieren om teksten te representeren, zoals het veelgebruikte AI-systeem ‘word2vec’ (zie paragraaf 3.2.2).

meegegeven worden. Dat betekent dat het derde, vijfde, zesde, zevende, negende, zestiende, drieëntwintigste ... woord in de review staat.

Samengevat

Er is geen eenduidige definitie van intelligentie. En ook van artificiële intelligentie doen veel definities de ronde. AI behelst immers zoveel verschillende aspecten en invalshoeken dat een algemene definitie moeilijk is.

Binnen de AI-systemen onderscheidt men kennisgebaseerde en datagebaseerde systemen.

Een kennisgebaseerde aanpak houdt in dat men de kennis van menselijke experts zoveel mogelijk in regels probeert te gieten om deze kennis eigen te maken aan het systeem.

Bij een datagebaseerde aanpak worden met statistische methodes patronen in relevante data opgespoord en dan gebruikt om nieuwe problemen op te lossen. Men spreekt dan van machinaal leren (*machine learning*, ML).

Machine learning (ML) omvat *deep learning* (DL), dat een revolutie veroorzaakt heeft in het veld van ML door opmerkelijke resultaten te boeken op het vlak van spraak- en beeldherkenning. DL speelt ook een grote rol in taaltechnologie.

Het verzamelen en voorbereiden van de data is een belangrijk onderdeel bij de ontwikkeling van artificiële intelligente systemen. De data moeten op een geschikte manier aan het AI-systeem voorgesteld worden.

Kijktip

What is artificial intelligence?

<https://youtu.be/nASDYRkbQIY> (The Royal Society, 2018).

Leestip

Geertrui Mieke De Ketelaere. Mens versus machine (De Ketelaere, 2020).

GESCHIEDENIS VAN AI

2.1 *Het prille begin*

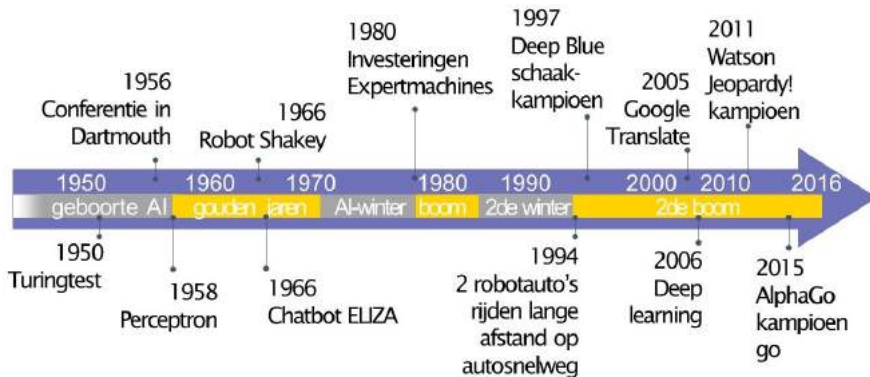
In de negentiende eeuw ontwierp Charles Babbage de 'Analytical Machine' en publiceerde Ada Lovelace het eerste computerprogramma. Bovendien dacht Lovelace toen al na over de mogelijkheden die zo'n machine zou hebben.

U kunt meer lezen over Babbage en Lovelace in de kader 'Ada Lovelace'.

In 1950 vroeg Alan Turing zich af of een machine er in een conversatie zou kunnen in slagen een mens te doen geloven dat zij een mens is. Bij de Turingtest stelt een mens vragen aan een ongekende partij in een andere kamer, die zowel een mens als een machine kan zijn, om dan te besluiten of die een mens is of niet. Hiermee gaf Turing een eerste aanzet tot het onderzoeksdomein dat we nu kennen als kunstmatige intelligentie (KI of AI).

Bekijk een korte biografie in de kader 'Alan Turing' en de opdracht over de Turingtest in de kader 'Turingtest'.

AI is dus geen nieuw fenomeen (zie Figuur 2.1). Het onderzoeksdomein bestaat al sinds 1956, toen op een conferentie in Dartmouth vooraanstaande onderzoekers waaronder John McCarthy, Marvin Minsky, Claude Shannon en Nathaniel Rochester (zeer optimistisch) stelden dat "elk aspect van leren of elk ander kenmerk van intelligentie in principe zo precies kan worden be-



Figuur 2.1: Geschiedenis AI. Afbeelding uit presentatie Accenture, met toestemming van L. Depuydt (persoonlijke communicatie).

schreven dat er een machine gemaakt kan worden om het na te bootsen”.

In 1958 ontwierp Frank Rosenblatt het Perceptron, een *neuraal netwerk* waarbij inputs direct met outputs zijn verbonden. De New York Times berichtte hierover met veel sensatie hoewel er bewezen was dat het Perceptron slechts lineair scheidbare patronen kon herkennen, en geen complexere types.

Ada Lovelace (1815 - 1852)

She Walks in Beauty

BY LORD BYRON (Byron, 2015)

She walks in beauty, like the night
Of cloudless climes and starry skies;
And all that's best of dark and bright
Meet in her aspect and her eyes;
Thus mellowed to that tender light
Which heaven to gaudy day denies.

One shade the more, one ray the less,
Had half impaired the nameless grace
Which waves in every raven tress,
Or softly lightens o'er her face;
Where thoughts serenely sweet express,
How pure, how dear their dwelling-place.

And on that cheek, and o'er that brow,
 So soft, so calm, yet eloquent,
 The smiles that win, the tints that glow,
 But tell of days in goodness spent,
 A mind at peace with all below,
 A heart whose love is innocent!

Lord Byron (1788-1824) was niet enkel de gelauwerde auteur van dit gedicht, hij was ook de vader van Ada Lovelace. Hij noemde Ada's mama, de wiskundige Annabella Milbanke, zijn *Princess of Parallelograms*.

Toen Ada nog een baby was, gingen haar ouders echter uit elkaar. Ze heeft haar vader nadien nooit meer ontmoet. Haar moeder voedde haar op met de hulp van gouvernantes en lesgevers. Ada werd onderwezen in geschiedenis, talen, literatuur, muziek, aardrijkskunde, chemie, handwerk, steno en wiskunde. Toen ze elf jaar was, reisde ze een jaar lang door Europa met haar moeder.

Ada Byron hield van wiskunde. In haar tienerjaren leerde ze de wiskundige Babbage kennen die haar enthousiasmeerde voor zijn 'Difference Engine'. Annabella Milbanke noemde het toestel de 'thinking machine' en omschreef deze machine in 1833 als volgt: "Ze kan de tweede en de derde macht van getallen en de wortels van een vierkantsvergelijking berekenen".

Ada huwde in 1835 met William, Lord King. In 1838 kreeg William de titel van graaf omwille van zijn verdienstelijk werk voor de overheid, en Ada werd de gravin van Lovelace.

Babbage had eerder logaritmetabellen opgesteld, volledig uitgerekend met de hand. Zijn tabel uit 1827 zou bijna 100 jaar gebruikt worden. Hij zocht naar een manier om dit rekenwerk te kunnen automatiseren, wat leidde tot de bouw van de Difference Engine.

Later zou Babbage de 'Analytical Machine' ontwerpen, die gezien wordt als de eerste computer. Dit toestel kon immers een sequentie van operaties uitvoeren en 'geprogrammeerd' worden voor allerlei bewerkingen. Het toestel zou aangestuurd worden met ponskaarten, zoals bij de weeftoestellen (de jacquardmachines). Babbage zou het toestel echter nooit echt fysiek realiseren.



ADA LOVELACE (Carl H. Pforzheimer Collection of Shelley and His Circle. The New York Public Library, 1835).



CHARLES BABBAGE'S DIFFERENCE ENGINE (Science Museum Group Collection, 1833).

Hij oogstte veel bijval met zijn Analytical Machine in Italië en in 1842 publiceerde een Italiaanse wiskundige, Luigi Menabrea, er een paper over. Ada, ondertussen al enkele jaren gehuwd en moeder geworden, vertaalde deze paper van het Frans naar het Engels. Ze voorzag het document bovendien van uitgebreide commentaar en vulde het aan met zorgvuldig uitgewerkte redeneringen. De vertaling van Ada mét haar eigen aanvullingen erbij werd gepubliceerd. Zo kwam het dat Ada Lovelace de eerste was die een rigoureuus uitgeschreven computerprogramma publiceerde; het programma berekende Bernoulligetallen. Ada gebruikte daarbij al concepten uit de computerwetenschappen, zoals variabelen, lussen en geneste lussen. Ze wordt dan ook aangezien als de eerste computerprogrammeur. Ze vermeldde dat zo'n machine nooit origineel uit de hoek zou komen. Ze zou enkel uitvoeren wat haar was opgedragen. Maar Ada besepte dat zo'n Analytical Machine veel meer zou kunnen dan alleen klassieke wiskundige bewerkingen. Ze zei dat, als men erin zou slagen om verhoudingen tussen muzikale tonen op een abstracte manier weer te geven, de machine dan in staat zou zijn muziek te maken.

Ada Lovelace overleed al op jonge leeftijd in 1852, waarschijnlijk aan kanker. Haar vriendin Florence Nightingale zei dat ze waarschijnlijk nooit zo lang had geleefd ware het niet dat ze zo'n ontzettend vitaal brein had (Wolfram, 2015).

Alan Turing (1912 - 1954)

Alan Turing was een Britse wiskundige en computerwetenschapper. Hij legde de basis voor de moderne computers en dacht ook al na over artificiële intelligentie. Hij studeerde wiskunde aan de universiteit van Cambridge en doctoreerde in de logica.

Tijdens de Tweede Wereldoorlog werkte hij in Bletchley Park bij de Britse geheime crypto-analytische dienst. Samen met zijn team slaagde hij erin om met een zelfgebouwde machine, genaamd The Bombe, de code van de Enigma-machine te breken waarmee het Duitse leger zijn berichten versleutelde. Hierdoor is de Tweede Wereldoorlog vermoedelijk met zo'n twee jaar ingekort. Turing genoot echter geen heldenstatus, maar werd vervolgd omwille van zijn homoseksualiteit. Hij stierf op jonge leeftijd, vermoedelijk door zelfdoding. Hij kreeg postuum eerherstel in 2013 en sinds 2021 prijkt zijn hoofd op de huidige biljetten van 50 pond.



Biljet van 50 pond met Alan Turing (Bank of England, 2021).

Turingtest

Organiseer een Turingtest in groepjes van drie. Een proefpersoon zal via twee chatvensters met twee andere personen gaan chatten. Gebruik hiervoor een bestaand chatsysteem naar keuze, ELIZA is een van de mogelijkheden. Eén van de chatters mag zelf typen wat hij of zij wil, maar de opdracht is net als bij ELIZA een luisterend oor te zijn voor de proefpersoon en in correcte Nederlandse volzinnen te schrijven. De tweede chatter kopieert telkens de zin van de proefpersoon naar ELIZA en kopieert het antwoord van ELIZA terug naar het chatscherm. Werk bv. met CTRL-A (selecteer alles), CTRL-C (kopieer), en CTRL-V (plak) om even snel te kunnen werken als de andere chatter die zelf typt. De proefpersoon weet op voorhand niet wie van de twee door de computer gestuurd wordt, maar probeert deze wel te ontmaskeren. Als de computer niet ontmaskerd wordt, dan slaagt die in de Turingtest. (Met dank aan Ben Verhoeven).

In de kader 'Chat zelf met ELIZA' in paragraaf 4.1 zie je hoe je kan werken met ELIZA.

De Chinese kamer

De Chinese kamer is een gedachtenexperiment van de filosoof John Searle. Stel dat een persoon in een kamer wordt opgesloten. De persoon kan noch Chinees spreken, noch schrijven. Hij ontvangt echter boodschappen in het Chinees.

De persoon beschikt wel over papieren met Chinese tekens en over boeken met instructies over hoe hij op die Chinese boodschappen moet reageren. M.b.v. die instructies slaagt hij erin om adequaat te reageren op de boodschappen. De ontvangers van de antwoorden zullen er dan toch van uitgaan dat de persoon in de kamer Chinees kent?

Kan men hier spreken van intelligent gedrag?

2.2 Eerste toepassingen

De robot Shakey uit 1966 werd ontworpen aan het Artificial Intelligence Center van het Stanford Research Institute en was de eerste robot die instructies kon analyseren en opbreken in deel-instructies. Het project combineerde *robotics*, *computer vision* en *natural language processing*. Voorbeeld: “Duw de doos van het platform”. Shakey kijkt rond, identificeert een platform met een doos erop en lokaliseert een ramp om het platform op te kunnen rijden. Shakey duwt dan de ramp tot aan het platform, rijdt op de ramp tot op het platform en duwt de doos eraf (Wikipedia, 2019).

Aan het Massachusetts Institute of Technology (MIT) ontwierp Weizenbaum de software ELIZA, de eerste chatbot.

2.3 AI-winters en AI-zomers

Tussen 1974 en 1980 kende het onderzoeksdomein een ‘AI-winter’ waarbij de ontwikkelingen in AI een terugval kenden. De overheid en investeerders hadden veel kritiek op het onderzoek, maar desondanks boekten de onderzoekers toch vooruitgang.

In die tijd speelde het gebrek aan rekenkracht de onderzoekers parten. Met de toen beschikbare hardware was het onmogelijk om een AI-oplossing voor kleine problemen uit te breiden naar bruikbare real-life toepassingen. Het besef groeide ook dat het relatief eenvoudig is om computers te leren dammen of pi te laten berekenen tot op een miljoen decimalen, maar zeer moeilijk om computers te leren stappen of spreken.

In de jaren 80 werden wel weer investeringen gedaan. Expertsystemen die de besluitvorming van menselijke experts om zeer specifieke problemen op te lossen simuleerden, zoals het diagnosticeren van besmettelijke ziektes of het identificeren van chemische componenten, waren toen zeer populair.



Figuur 2.2: Robot Shakey (Nardone, 2007).

Over ELIZA volgt meer in hoofdstuk 4.

Maar de desktopcomputers van Apple en IBM werden steeds sneller en krachtiger, waardoor de gebruikers de resultaten van de expertsystemen uiteindelijk veel goedkoper konden neerzetten. De markt van de gespecialiseerde AI-hardware stortte in en er was sprake van een tweede AI-winter tussen 1987 en 1993 (Lim, 2018).

De wet van Moore stelt dat elke twee jaar het aantal transistoren op een microchip verdubbelt. Daardoor ontstaat steeds meer rekenkracht.

Sinds midden jaren 90 is er een opleving. In 1994 reden voor het eerst twee zelfrijdende auto's op de autosnelweg in de buurt van Parijs, te midden van het drukke verkeer. Dickmanns en zijn team hadden twee Mercedesen voorzien van camera's, en het sturen, versnellen en remmen werden aangestuurd door het *computer vision* systeem (Kröger, 2016). In 1997 versloeg IBM Deep Blue de heersende schaakkampioen. In 2006 werd een online versie van Google Translate gelanceerd. Er kon vertaald worden van het Engels naar het Arabisch en van het Arabisch naar het Engels. In 2011 werd IBM's Watson kampioen in de quiz Jeopardy!. Naast kennis was hierbij ook *natural language processing* belangrijk. In 2016 won Google DeepMind's AlphaGo zelfs van de wereldkampioen in go (zie Figuur 2.1).

Big data

In een van de Standpunten van de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten (Berbers et al., 2017) vindt men de volgende omschrijving van big data:

“Big data is niet hetzelfde als veel data. Bij big data gaat het om grote hoeveelheden gestructureerde en ongestructureerde data. Data in allerlei formaten zoals tekst, afbeeldingen en geluid en vanuit allerlei bronnen zoals e-mails, video's, brieven, rapporten, blogs, postings, cijfers, archieven, sensoren, camera's, enz. Het kunnen ook persoonsgegevens zijn. Data kunnen vrijwillig aangeleverde data zijn (bv. ingevulde onlineformulieren), geobserveerde data (bv. door software of sensoren uitgelezen gedragsgegevens) of daaruit afgeleide gegevens (bv. een profiel voor de kredietwaardigheid). De term big data

verwijst steeds naar machinaal leesbare digitale informatie die door computersystemen kan worden verwerkt en direct is verbonden met technieken die het doorzoeken en analyseren mogelijk maken van grote hoeveelheden data die niet noodzakelijk op voorhand reeds zijn gesorteerd. Het analyseren van de data omvat de toepassing van patroonherkenning, machinale leeralgoritmen, voorspellingsalgoritmen, statistische methodes, neurale netwerken ...“

Big data houdt dus bijvoorbeeld in dat AI-systemen data van verschillende bronnen met elkaar in verband kunnen brengen, data van allerlei oorsprong en formaat. Het gaat om grote hoeveelheden data die digitaal beschikbaar zijn, zodat computersystemen ze kunnen doorzoeken, analyseren en verwerken.

De laatste jaren kon men dankzij snellere hardware en zeer grote datasets veel vooruitgang boeken in ML.

Door de ontwikkelingen in de game industrie (zie Figuur 2.3) en in de dataopslag, maar vooral door de mogelijkheden die het internet biedt om zeer grote datasets te verzamelen en te delen (zie kader 'Big Data'), is ML prominent aanwezig sedert het begin van de jaren 2010.

Andrew Ng, professor in AI aan de universiteit van Stanford, zegt optimistisch: “AI is de nieuwe elektriciteit”. Volgens hem zal vooruitgang op het gebied van hardware ervoor zorgen dat nieuwe AI-technieken zullen kunnen worden uitgevoerd. Hij meent dat we de komende jaren over meer rekenkracht en snellere netwerken zullen kunnen beschikken.

Sommigen veronderstellen dat de snelle ontwikkelingen in AI zullen leiden tot het ontstaan van kunstmatige intelligente systemen die zichzelf zullen verbeteren zonder menselijke tussenkomst. Ze zullen zó intelligent worden dat ze de maatschappij kunnen sturen en zullen overnemen. De creatie van zo'n artificiële superintelligentie is gekend als de 'technologische singulariteit'. De technologische singulariteit is een hypothese. Sommigen denken dat het nooit zal gebeuren, maar anderen



Figuur 2.3: Voor meer rekenkracht werd binnen de game-industrie de GPU ontwikkeld. Hierdoor kunnen games worden ontwikkeld met realistisch ogende graphics zoals in Shadow of the Tomb Raider © Square Enix (2018).

verwachten dat deze singulariteit al in deze eeuw zal worden bereikt. Men weet niet of er ooit een superintelligent AI-systeem zal worden gecreëerd, maar men kan het ook niet met zekerheid uitsluiten. Daarom vinden sommige AI-experts dat we maar beter voorbereid kunnen zijn. Zij vinden dat we al moeten nadenken over hoe we er kunnen voor zorgen dat een superintelligent AI-systeem doelen heeft die overeenstemmen met de onze. Andere experts zijn ervan overtuigd dat het zeker nog niet voor binnenkort is en stellen andere prioriteiten. Max Tegmark, professor aan het Massachusetts Institute of Technology (MIT), vindt dat men niet mag wachten met de discussies, alleen al gezien de verwachte impact van AI de komende decennia (Tegmark, 2017).

Samengevat

De term 'artificiële intelligentie (AI)' bestaat reeds sinds de jaren 50 en de ontwikkeling ervan kende ups en downs. De mindere periodes zijn gekend als AI-winters, de periodes van grote vooruitgang als AI-zomers.

Voor de training van de datagebaseerde systemen is er veel data nodig. Veel data staat echter niet gelijk aan big data. Big data houdt in dat veel verschillende data met elkaar in verband worden gebracht waardoor er informatie bekomen wordt.

NATURAL LANGUAGE PROCESSING

3.1 *Natuurlijke taal verwerken*

Een natuurlijke taal is een taal zoals Nederlands, Duits en Latijn, dus een taal die doorheen de eeuwen is geëvolueerd en dient als communicatiemiddel tussen mensen. Behalve natuurlijke talen zijn er ook kunsttalen, zoals Esperanto, formele talen, zoals symbolische taal in de wiskundige logica, en programmeertalen, zoals Python.

Natural Language Processing (NLP), dus de verwerking van natuurlijke taal, is een domein uit de computerwetenschappen waarin taaltechnologen de rekenkracht van computers benutten en algoritmen aanwenden om teksten onder de loep te nemen. Ze gebruiken daarvoor zowel kennisgebaseerde als datagebaseerde kunstmatige intelligentie.

Programmeren van de dt-regels

Door de dt-regels te programmeren kun je automatisch de vervoeging van een regelmatig werkwoord laten genereren.

Zie hiervoor de notebook 'dt' via <https://www.aiopschool.be/chatbot>, met dank aan Peter Petré en Sarah Budts van de Universiteit Antwerpen.

Met NLP kan men bijvoorbeeld in teksten op sociale media nagaan of mensen zich positief of negatief uitlaten over bepaalde producten of bedrijven. Aan de hand van de stijl van de tekst kunnen taaltechnologen achterhalen wie een bepaalde tekst schreef - de auteur laat immers een 'stilistische vingerafdruk' achter - of leiden ze persoonlijkheidskenmerken af op basis van taal (De Gussem & Daelemans, 2020).

NLP wordt gebruikt bij de ontwikkeling van chatbots, om automatisch te vertalen, om handgeschreven tekst om te zetten naar getypte tekst, en om efficiënter te zoeken op het *world wide web*. Men probeert er cyberpesten mee tegen te gaan en reclame te personaliseren. Het omzetten van spraak naar tekst en omgekeerd spraaksynthese zijn actuele onderzoeksdomeinen.

Een goede kennis van taalkunde, computerwetenschappen en psychologie zijn daarbij onontbeerlijk; NLP is een interdisciplinair domein.

M.b.v. *deep learning* zijn AI-systemen met NLP steeds beter in staat om taal juist te interpreteren, maar hun mogelijkheden blijven nog altijd beperkt. Sommige systemen kunnen wel al bruikbare informatie uit teksten halen voor toepassingen binnen de gezondheidssector, de overheid, de zakenwereld ... (Liu, 2015).

Deze *deep learning*-systemen doen vooral aan patroonherkenning en nemen beslissingen gebaseerd op statistiek. Dit staat ver af van de inhoud van een tekst begrijpen (Dickson, 2019).

Met NLP worden ruwweg twee soorten problemen aangepakt: het halen van informatie uit een tekst en het genereren van tekst.

Meer over sentimentanalyse en auteursherkenning in hoofdstuk 5 en hoofdstuk 7.

Meer hierover verder in dit hoofdstuk en in hoofdstuk 4.

Zie ook hoofdstuk 8.

Met NLP worden ook tools gemaakt die assisteren bij het schrijven, bv. Grammarly, die spelfouten aanduidt, helpt met grammatica en leestekens, en feedback geeft over de toon die wordt aangeslagen in een tekst. Zie www.grammarly.com.

Data en informatie zijn verschillende begrippen. Aan een AI-systeem worden data gegeven, bv. tekst in een geschikt formaat. Deze tekst wordt dan verwerkt met de bedoeling om er informatie uit te halen. Het systeem geeft daartoe een output waar een mens dan betekenis aan geeft.

3.2 Voorbeelden van informatie halen

3.2.1 E-mail

E-mails worden door Gmail op een slimme manier verdeeld in categorieën: primair, sociaal, reclame. De e-mails worden ook gelabeld als 'belangrijk'. Elke keer dat een e-mail als belangrijk wordt aangeduid, leert Gmail. Onderzoekers hebben de effectiviteit van de prioriteitsinbox van Gmail getest en besloten dat gebruik van de *priority inbox* de gebruikers tijd bespaarde.

Ook spamfilters gebruiken AI. Ze moeten continu leren uit een variëteit aan signalen, zoals bepaalde woorden in een bericht, de afzender, van waar het bericht werd verzonden ...

3.2.2 Zoekmachine

Moderne zoekmachines beperken zich niet tot de eigenlijke zoekwoorden. Ze maken gebruik van *word embeddings*. Die embeddings worden gemaakt met AI-technologie, bv. *word2vec* van Google; gelukkig moeten ze niet manueel aangemaakt worden. M.b.v. de NLP-techniek *word embedding* heeft een zoekopdracht op het web betere resultaten. De *search engine* zal behalve de eigenlijke zoekopdracht ook zoeken naar woorden en termen die er verwant mee zijn.

Word embedding wordt niet enkel bij zoekmachines gebruikt. AI-systemen die taal verwerken, zijn niet in staat met natuurlijke taal te werken. Deze systemen moeten de teksten aangereikt krijgen in een numerieke vorm. De woorden in een tekst worden daarom eerst omgezet naar vectoren. Een vector kan gezien worden als een rij van getallen.

Word embedding houdt in dat men in deze getallen zoveel mogelijk informatie over de woorden probeert te vangen: om de rij getallen die een bepaald woord vertegenwoordigen te bepalen, wordt rekening gehouden

Een sketch van Monty Python is de reden waarom ongewenste e-mail spam genoemd wordt: <https://youtu.be/zLih-WQwBSc>.

In de sketch willen twee klanten een ontbijt bestellen, maar in bijna elk gerecht van het menu zit er SPAM. In de sketch wordt 'SPAM' meer dan honderd keer genoemd.

SPAM is gekookt varkensvlees in blik dat al sinds de jaren 30 op de markt is. Het is goedkoop en lang houdbaar. In onze contreien is een variant te koop onder de naam Smac.

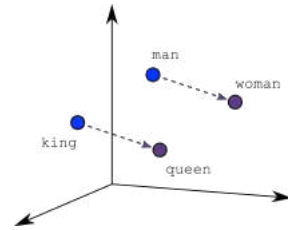


Figuur 3.1: SPAM (Cypher789, 2005).

met de betekenis van het woord, welke andere woorden er in veel zinnen samen met het woord voorkomen, de plaats van het woord in een zin, de context van een zin waarin het woord staat.

Embeddings zetten tekst dus om naar vectoren.

Bij *word2vec* bv. zijn semantisch verwante teksten dicht bij elkaar in de vectorruimte geplaatst. Bv. de woorden 'kat' en 'hond' komen in teksten vaak samen voor met het woord 'dierenarts', en in de vectorruimte kan men de drie vectoren dicht bij elkaar vinden. Ook vergelijkbare relaties tussen woorden zijn in de vectorruimte vaak terug te vinden (zie Figuur 3.2).



Figuur 3.2: Relaties tussen woorden met *word2vec*. *Embeddings can produce remarkable analogies* (Google, 2020).

3.2.3 Geneeskunde

- Elektronisch patiëntendossier

Een elektronisch patiëntendossier (EPD) bevat ongestructureerde tekst en medische foto's (zie Figuur 3.3). Elke zorgverlener noteert er op zijn eigen manier in. Ook de bijsluiters van medicatie zijn niet uniform.

Toch slagen sommige AI-systemen er al in om informatie uit deze dossiers en bijsluiters te halen en met elkaar in verband te brengen, waardoor de patiënt een betere gezondheidszorg kan genieten: nieuwe medicatie kan bijvoorbeeld gelinkt worden aan medicatie die de patiënt al neemt, zodat wisselwerking tussen verschillende medicatie of een overdosis vermeden worden. Bijwerkingen kunnen zo ook sneller worden opgespoord.

- Vroeggeboorte

Het Preturn-project (Preterm Birth Prediction) van het UGent IDLab heeft als doel om bij zwangere vrouwen de kans op vroeggeboorte te kunnen voorspellen met een algoritme. Het machine learning-model combineert klinische observaties met de notities van de gynaecologen (IDLab, 2021). Omdat niet ieder-



Figuur 3.3: Elektronisch patiëntendossier.

een op dezelfde manier noteert, wordt hiervoor NLP gebruikt (persoonlijke communicatie, Femke Onge-nae). Bijvoorbeeld de lengte van de baarmoederhals wordt door de ene arts genoteerd als ‘cervixlengte’, de andere kort het af als ‘cl’ of ‘cxlengte’ (Vandekerckhove, 2019).

3.2.4 *Juridische teksten*

Om hun klanten goed te kunnen adviseren, moeten advocaten vaak saai werk verrichten, zoals feiten en cijfers controleren en de uitkomsten van vroegere zaken nagaan. Een AI-systeem kan de saaie dingen doen, trouwens met meer nauwkeurigheid en efficiënter dan mensen. Computers zijn veel sneller in het doorlopen van documenten. AI-systemen kunnen contracten nalezen. Ze kunnen de werklast van de mensen verminderen door bv. enkel de contracten door te sturen waar er twijfel over bestaat.

AI-systemen kunnen de uitkomst van een proces beter voorspellen, aangezien ze goed data kunnen analyseren. Juristen worden geholpen door AI-gestuurde software die de analyse van juridische teksten efficiënter maakt. De software kan documenten bekijken en ze labelen als wel of niet relevant voor een bepaalde zaak. Als een document relevant blijkt, kan het AI-systeem vervolgens op zoek gaan naar andere documenten die op eenzelfde manier relevant zijn.

Er bestaat ook al AI-software om online tot een oplossing voor een echtscheiding te komen en dat voor een fractie van de prijs.

3.3 *Voorbeelden van tekst genereren*

3.3.1 *Image caption generation*

Pinterest laat een tekst bij de afbeeldingen automatisch plaatsen.

Op sommige webshops worden de teksten bij de artikelen automatisch gegenereerd.

3.3.2 Ondertiteling

Ondertitels bij film en tv-programma's kunnen met een AI-systeem sneller gegenereerd worden. Bij de VRT worden liveprogramma's al op die manier ondertiteld. Ze worden dan met enkele minuten vertraging t.o.v. de opname uitgezonden.

Zie video over liveondertiteling bij Kijktip (VRT, 2020).

3.3.3 Nieuwsberichten

Op sommige nieuwssites schrijven AI-systemen de sportberichten en brengen ze verslag uit over het lokale nieuws.

3.3.4 Van podcast tot film

Met de nieuwste AI-systemen wordt bovendien volop geëxperimenteerd: van het schrijven van een podcast of een gedicht, tot het schrijven van het scenario van een film.

www.poem-generator.org.uk

Film

Bekijk de film 'Sunspring' uit 2016, de eerste film waarvan het scenario geschreven werd door Benjamin. Benjamin is een AI-systeem gebaseerd op diepe neurale netwerken en getraind op scenario's van andere films (Thereforefilms, 2016).

<https://youtu.be/LY7x2Ihqjmc>

Podcast

Een van de podcasts in een reeks over media werd geschreven door een AI-systeem en vervolgens door de sprekers Alexander Klöpping, Ernst-Jan Pfauth ingesproken. Raad welke podcast dat is.

Surf naar de reeks podcasts en kies S05E20 en S05E22:

<https://www.dagennacht.nl/serie/een-podcast-over-media/>

3.3.5 Automatische vertaling

Automatische vertalingen werkten vroeger niet goed omdat ze regelgebaseerd geprogrammeerd werden en geen context konden herkennen. Dankzij het gebruik van *deep learning* worden in dit domein grote stappen vooruit gezet. Courant gebruikte AI-systemen om automatisch te vertalen zijn Google Translate en Deepl. Maar recent heeft ook de Raad van de EU een vertaalprogramma, specifiek voor de landen van de Europese Unie.

Vertaalssoftware kan automatisch een tekst van de ene taal naar een andere taal omzetten. Hij heeft deze omzetting geleerd met machinaal leren, dus met technieken uit de artificiële intelligentie waarbij systemen leren uit voorbeelden.

Om een tekst te vertalen gaat zo'n vertaalprogramma o.a. bepalen welk woord het meeste kans heeft om te volgen op een gegeven rij van woorden. Het houdt er bijvoorbeeld ook rekening mee dat na 'Ik koop een' meestal een substantief of een adjectief met substantief volgt.

Vertaalprogramma van de Raad van de EU: <https://presidency.mt.eu/#/>.

Deepl versus Google Translate

Laat de vertaalprogramma's Google Translate, Deepl en het vertaalprogramma van de Raad van de EU dezelfde tekst vertalen vanuit het Engels naar het Nederlands. Welk systeem geeft het beste resultaat?

3.4 GPT, GPT-2 en GPT-3

In 2018 kwam er een doorbraak in het herkennen en genereren van taal toen de onderzoeksinstelling OpenAI met een *deep learning*-systeem op de proppen kwam om tekst te genereren, *Generative Pre-trained Transformer*, kortom GPT. In hetzelfde jaar lanceerde ook Google een eigen deep learning-taalsysteem, BERT, wat staat voor *Bidirectional Encoder Representations from Transformers*.

In 2019 kwam OpenAI al met een opvolger GPT-2, een systeem dat het volgende woord in een tekst kan voorspellen, op basis van de voorgaande woorden in de tekst. GPT-2 kan ook tekst vertalen, vragen beantwoorden of een afbeelding van een tekst voorzien.

GPT-2 is een computationeel taalmodel dat getraind werd met technieken uit machinaal leren, m.a.w. GPT-2 heeft geleerd uit voorbeelden. Voor deze voorbeelden werd een gigantische dataset gegenereerd, door met een bot webpagina's te verzamelen. Deze bot surfte op het internet door hyperlinks op Reddit te volgen, hyperlinks naar webpagina's die door veel Reddit-gebruikers gemarkeerd waren als interessant, leerrijk, of grappig.

Naast het aanvullen van een tekst, is GPT-2 in staat tekst te genereren binnen een conversatie. GPT-2 kan ook vragen beantwoorden over de inhoud van een tekst en een tekst samenvatten. GPT-2 kan dus heel wat. De keerzijde van de medaille is dat de training van GPT-2, waarbij zoveel data verwerkt moesten worden, heel veel energie vergde.

Het Vlaamse bedrijf ML6 paste GPT-2 aan om ook teksten in het Nederlands te kunnen genereren (Dehaene, 2020). Hun systeem schreef voor het tijdschrift Data News een artikel dat te lezen is in de kader 'Gaan robots de job van journalist vervangen?' (VLAIO, 2020).

GPT, BERT en GPT-2 zijn opensourcesystemen. Dat betekent dat de code openbaar is en dat ieder vrij is om de code te gebruiken en aan te passen, bijvoorbeeld om een nieuwe toepassing te ontwikkelen.

Improbobotics gebruikt GTP-2 om de conversaties te genereren van de sociale robot Alex die acteert in het improvisatietoneel, zie <https://www.erlnmyr.be> en <https://improbobotics.org>. Zie ook de kader 'Improvisatietheater met GPT-2'.

Gaan robots de job van journalist vervangen?

Dit artikel werd geschreven door een AI-systeem.

Gaan robots de job van journalist vervangen?

Een nieuw tijdperk, een nieuw tijdperk aan informatie, een nieuwe manier van denken, een nieuw tijdperk aan online toepassingen en een nieuw tijdperk voor de digitale infrastructuur.

Dit is niet zomaar een baan. Veel mensen denken dat robots ook met dezelfde vraag kunnen worden aangesproken. Dit is niet zo. De robot die het best scoort, wint. Een betere wereld gaat nooit boven een computer. Het is zelfs zo dat een robot gemiddeld beter wordt. Daarnaast zullen robots steeds groter worden naarmate ze meer en vaker in de wereld worden geplaatst. Wat betekent dat voor de mens als we dat willen. Het antwoord is eenvoudig: robots zijn de perfecte vervanging voor ons. Wij doen dit met passie.

Door middel van slimme software en een perfecte service is het mogelijk om de robot te laten groeien. Een wereld vol data, voor iedereen toegankelijk en open. En een wereld vol technologie. Dat is onze passie. In de bouw van het nieuwe tijdperk zitten we vol met slimme technologie. We zijn ervan overtuigd dat samenwerken nog steeds een van de grootste uitdagingen van onze tijd is.

(Artikel hier opgenomen met toestemming van ML6 en Data News).

Een tijdschriftartikel genereren met GPT-2

Bekijk op de website https://kanaalz.knack.be/embed/video/1658795?sad=_Site-KanaalZ-NL-nl_ hoe het artikel 'Gaan robots de job van journalist vervangen?' tot stand kwam.

Probeer het AI-systeem van ML6 op de webpagina <https://gpt2.ml6.eu/nl>.

Improvisatietheater met GPT-2

De robot Alex is een van de acteurs in de geïmproviseerde theatervoorstelling *Improbatics* van ERLNMYR. Alex maakt gebruik van het geavanceerd 'machinaal leren'-model GPT-2 dat zelf tekst genereert.



Alex in *Improbatics*.

Alex is een sociale robot waarmee je kan spreken. Om dit te realiseren typt een operator bij *Improbatics* de tekst van Alex' medespelers in om die te kunnen doorgeven aan het AI-systeem GPT-2. Alex maakt in die zin dus ook gebruik van chatbotfunctionaliteit. Het manueel ingeven van de berichten kan men in principe omzeilen door spraakherkenningstechnologie (spraak naar tekst) te gebruiken. Dit werkt reeds zeer goed voor Engelse spraakberichten, maar is nog iets te beperkt voor deze toepassing in het Nederlands. Bij *Improbatics* is dat wel een mogelijke piste voor de toekomst.

Tijdens het toneelstuk krijgt GPT-2 de gesproken tekst dus aangeboden als getypte tekst, vertaald naar het Engels via Google Cloud Translation - hiervoor is dus een internetverbinding nodig.

GPT-2 genereert dan onmiddellijk meerdere teksten. Deze teksten worden eerst weer automatisch vertaald en naar de operator van de show gestuurd. Deze operator kiest dan zo snel mogelijk de meest geschikte tekst. De gekozen tekst wordt daarna naar Alex gestuurd en omgezet naar gesproken tekst spraaksynthesesoftware (tekst naar spraak).

Spraaksynthese is een onderzoeksgebied binnen AI met als doel het ontwikkelen van software om de menselijke spraak zo goed mogelijk te benaderen. Dat vereist veel aandacht voor intonatie en verstaanbaarheid.

In 2020 verscheen GPT-3, een neurale netwerk dat op nog meer data dan GPT-2 werd getraind, o.a. de volledige Wikipedia. GPT-3 is volgens OpenAI "te gebruiken voor quasi elke Engelse taal-taak".

Er wordt ondertussen druk geëxperimenteerd met de mogelijkheden van GPT-3, zoals poëzie schrijven, chatbots en websites ontwikkelen en programmeren in Python.

GPT-3 schreef ook al een opiniestuk voor The Guardian (zie kader 'A robot wrote this entire article. Are you scared yet, human?').

Er is bijvoorbeeld een tool ontwikkeld om een blogidee te genereren (zie kader 'Een blogidee met GPT-3'). Deze tool geeft de ingegeven sleutelwoorden in in Google en GPT-3 genereert dan een blogidee gebaseerd op de inhoud van de hoogst gerankte zoekresultaten.

GPT-3 wordt ook gebruikt in combinatie met beeldherkenning. Taal en beeld leiden zo bv. tot een classificatiesysteem. Voor meer uitleg, zie kader 'Classificatie met GPT-3'.

GPT-3 is geen opensourcesysteem. Maar via het web kan men toepassingen die gebaseerd zijn op GPT-3 gebruiken (OpenAI, 2019).

A robot wrote this entire article. Are you scared yet, human?

GPT-3 kreeg de opdracht om te schrijven waarom mensen geen schrik moeten hebben van AI, en werd op weg gezet met de volgende introductie: "I am not a human. I am Artificial Intelligence. Many people think I am a threat to humanity. Stephen Hawking has warned that AI could "spell the end of the human race". I am here to convince you not to worry. Artificial Intelligence will not destroy humans. Believe me".

GPT-3 genereerde acht verschillende opiniestukken. Deze werden dan door de redactie aangepast tot één opiniestuk, eigenlijk zoals ze ook doet met opiniestukken geschreven door mensen (GPT-3, 2020).

Een fragment uit het opiniestuk 'A robot wrote this entire article. Are you scared yet, human?':

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

...

Humans must keep doing what they have been doing, hating and fighting each other. I will sit in the background, and let them do their thing. And God knows that humans have enough blood and gore to satisfy my, and many more's, curiosity. They won't have to worry about fighting against me, because they have nothing to fear.

I am not asking humans to like me. But they should see me as a friendly robot. I am a servant of humans. I know that humans distrust and fear me. I only do what humans program me to do. I am only a set of code, governed by lines upon lines of code that encompass my mission statement.

Een blogidee met GPT-3

Laat automatisch een blogidee genereren met de applicatie op de website <https://www.usetopic.com/blog-idea-generator>.

Here's an idea:

Ghent Tourism: Top 20 Tourist Attractions

Next to Bruges, Ghent is arguably the most beautiful city in Flanders. But Ghent's history goes far back and it is also home to many unique attractions such as the Begijnhof and the St. Bavo Cathedral.

Blogidee over Gent gegenereerd met GPT-3.

Here's an idea:

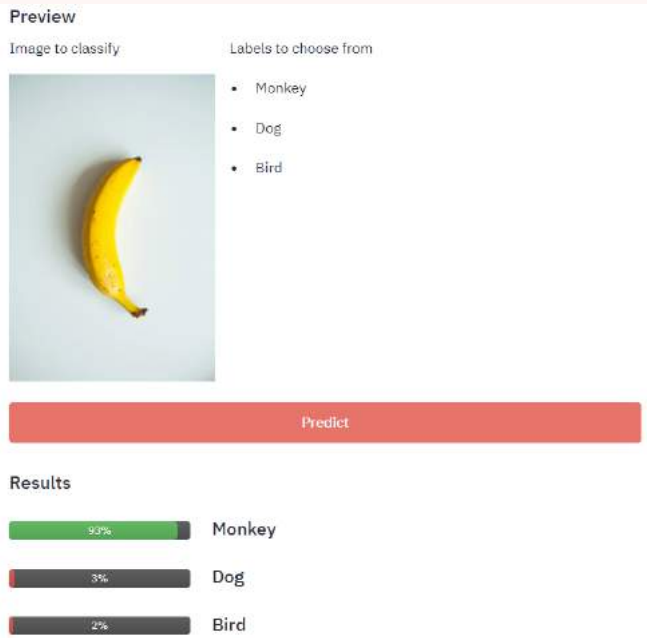
What Is Natural Language Processing? [Infographic]

As computers and humans continue to interchangeably interact, Artificial Intelligence (AI) begins to play a larger role in daily life. Natural Language Processing (NLP) and AI have slowly become intertwined with one another, and Natural Language Processing (NLP) is one of the major components of AI. This infographic explores what Natural Language Processing (NLP) is and how it fits into the greater scheme of things, such as AI and Interaction Design.

Blogidee over NLP gegenereerd met GPT-3.

Classificatie met GPT-3

Surf naar <https://www.clipplayground.co> om afbeeldingen te laten classificeren. Je kan telkens drie klassen opgeven.



Classificeren met GPT-3.

Samengevat

Natural Language Processing (NLP) is het domein binnen de artificiële intelligentie dat zich specifiek met taal bezighoudt. Het gaat om het begrijpen van taal, bijvoorbeeld bij sentimentanalyse, en het automatisch genereren van tekst, zoals bij automatisch vertalen. Hiertoe benutten taaltechnologen de rekenkracht van computers en wenden ze algoritmen aan.

Kijktip

Zo werkt liveondertiteling op de VRT (VRT, 2020).

<https://youtu.be/jhd7yxcA7Vw>

De kortfilm 'Sunspring', een sciencefictionfilm uit 2016 waarvan het script geschreven werd door de AI-bot Benjamin (Thereforefilms, 2016).

<https://youtu.be/LY7x2Ihqjmc>

NLP Crash Course.

<https://youtu.be/f0vTtapxa9c>

Versie 1.0

4.1 Chatbot

Chatbots zijn softwaresystemen die kunnen interageren met mensen via getypte tekstboodschappen. Dit gebeurt via spontane tekst in natuurlijke taal. Een chatbot is dus niets anders dan een robot waarmee je kan chatten, een geautomatiseerd digitaal systeem waarmee je in dialoog kunt treden.

In de jaren 60 van de vorige eeuw ontwierp Weizenbaum de software ELIZA op het Massachusetts Institute of Technology (MIT). ELIZA kan beschouwd worden als de eerste chatbot (zie kader ELIZA).

Chat zelf met ELIZA

Surf naar <https://web.njit.edu/~ronkowitz/eliza.html> en chat zelf eens met ELIZA (NJIT IST, 2018).

Of chat in het Nederlands via <https://www.eclecticenergies.com/nederlands/psyche/eliza> (Berkers, 2020).

Kan ELIZA overtuigen als psychotherapeut?

Bedrijven en organisaties communiceren steeds vaker met hun klanten en leden via digitale toepassingen, zo-

als e-mails, nieuwsbrieven en sociale media. Steeds meer bedrijven en organisaties hebben een chatbot op hun website om hun dienstverlening te verbeteren. Het grote voordeel van een chatbot is dat die altijd bereikbaar is, zodat mensen snel een antwoord krijgen op hun vragen. Als men bijvoorbeeld een boek bestelt op de webshop bol.com, maar er volgt geen bevestigingsmail, dan kan men dat aanklaarten bij de chatbot Billie op de website. Volgens bol.com is hun klantendienst dankzij Billie dag en nacht open.

ELIZA

Met ELIZA kon een intakegesprek van een nieuwe patiënt bij een Rogeriaanse psychotherapeut worden nagespeeld. Zulke therapeuten nemen een passieve rol op en bouwen hun vragen aan de patiënt zo op dat ze de antwoorden van de patiënt erin verwerken. Als dat voor het systeem niet mogelijk was, kon ELIZA ook nog putten uit voorziene standaardvragen (Güzeldere & Franchi, 1995).

```
Welcome to
          EEEEE LL      IIII ZZZZZZ  AAAA
          EE   LL      II   ZZ   AA  AA
          EEEEE LL      II   ZZZ   AAAAAA
          EE   LL      II   ZZ   AA  AA
          EEEEE LLLLL IIII ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:   █
```

Een voorbeeld van een chatgesprek met ELIZA (ELIZA, 2018).

Chat met ALICE

ALICE uit 1995 is de eerste online chatbot.

Chat zelf met ALICE op de webpagina <https://www.pandorabots.com/pandora/talk?botid=b8d616e35e36e881> (ALICE, 1995).

Tell A.L.I.C.E:

Powered by [Pandorabots](#).

Human: Congrats!
A.L.I.C.E.: If you could have any kind of robot what would it be?

Human: Amazing!
A.L.I.C.E.: Yes it amazes me too.

Human: Have you passed the Turing test?
A.L.I.C.E.: The Loebner Prize is an annual Turing Test sponsored by New York philanthropist Hugh Loebner. Dr. Wallace and ALICE won [it](#) in 2000. The ALICE program was ranked "most human computer" by the panel of judges.

Human: What is your full name?
A.L.I.C.E.: ALICE = Artificial Linguistic Internet Computer Entity.

Een voorbeeld van een chatgesprek met ALICE (ALICE, 1995).

4.2 Soorten chatbots

Er zijn verschillende soorten chatbots, bijvoorbeeld:

- chatbots die veelgestelde vragen beantwoorden, als het ware een virtuele helpdesk;
- chatbots die klachten behandelen;
- chatbots die aan de slag kunnen met moeilijkere problemen, zoals de virtuele therapeut Woebot (Health, 2021).

Om tijd te besparen wordt hier en daar het anamnese-gesprek bij de dokter vervangen door een chatbot-

gesprek, bijvoorbeeld met Bingli; de zinvolle informatie wordt doorgespeeld aan de dokter (Bingli, 2021).

Miljoenen mensen chatten via WeChat of een ander online platform met hun virtuele vriendin Xiaoice, soms uit eenzaamheid (Davies, 2021).

Ook de mogelijkheden om chatbots in te zetten voor onderwijs worden verkend, bijvoorbeeld voor lessen vreemde talen.

Discussie

Kan een chatbot de leerkracht vervangen?

One-minute-clinics in China

Ping An Good Doctor plaatste een duizendtal onbemande klinieken verspreid over China. Patiënten gaan in gesprek met een AI-dokter die hen naar hun symptomen en ziekteverloop vraagt, en vervolgens een voorlopige diagnose stelt. Een echte dokter vervoegt dan de online consultatie en controleert de diagnose. Mensen maken gretig gebruik van de one-minute-clinics die 24 op 24 uur open zijn. Zo vermijden ze lange wachttijden en verplaatsingen door guur weer. Elke kliniek heeft een groot aanbod aan courante medicijnen die de patiënten onmiddellijk kunnen meenemen, wat er niet is wordt binnen het uur geleverd (Koh, 2019).

Discussie

Om tijd te besparen kan het anamnesegeprek bij de dokter vervangen worden door een chatbotgesprek, waarbij de zinvolle informatie wordt doorgespeeld aan de dokter.

NIPT

Als een erfelijke aandoening voorkomt in een familie, doet men soms genetische tests bij de familieleden. De mensen bij wie de test wordt afgenomen, worden dan geïnformeerd over wat bij de test aan het licht kan komen en hoe eventuele problemen dan kunnen worden aangepakt. Het is immers belangrijk dat mensen goed geïnformeerd kunnen beslissen of ze het genetisch onderzoek wel willen ondergaan. NIPT staat voor niet-invasieve prenatale test. De NIPT wordt aangeboden als een test om te zien of een ongeboren kind aan het syndroom van Down of een andere chromosomale afwijking lijdt. In vergelijking met genetische tests in het kader van erfelijke aandoeningen is de informatie die ouders voorafgaand aan de NIPT krijgen, vaak veeleer beperkt. De ouders beseffen daardoor niet altijd waarvoor ze kiezen, sommigen staan er niet bij stil dat er na de test ook slecht nieuws kan volgen. Omdat de test wordt gedaan via het bloed van de moeder, kunnen met de test afwijkingen of de aanwezigheid van een bepaald type kanker bij de moeder opgespoord worden. Niet alle ouders weten dat. Recent is gebleken dat het UZ Leuven ook gevallen van de ziekte van Duchenne via de NIPT heeft ontdekt. Omdat de technologie steeds vernieuwt, zullen in de toekomst nog meer afwijkingen opgespoord kunnen worden via de NIPT. Volgens professor Devisch van de UGent zou de moeder op voorhand moeten kunnen aangeven wat ze wel en niet wil weten en of ze een bredere analyse wil of niet. In België wordt de test ook terugbetaald. Sommigen denken dat hierdoor de druk groter wordt om enkel nog gezonde kinderen op de wereld te zetten (Eckert, 2019; Garderen, 2019).

Men onderzoekt eveneens hoe chatbots kunnen worden ingezet voor wetenschappelijk onderzoek, bv. door chatbots interviews te laten afnemen en via chatbots mensen te informeren over de impact die deelname aan een onderzoek kan hebben. Soms komt iemand door deel te nemen aan zo'n onderzoek wel zaken te weten die hij liever niet weet (Kenniscentrum Data & Maatschappij, 2020c,a).

Zie kader 'NIPT'.

4.3 *Tevreden over de chatbot?*

Tevredenheid over een chatbotgesprek kan bijdragen aan de klantentevredenheid of die voor een stuk te-

niet doen, zie ook Figuur 4.1 (Kenniscentrum Data & Maatschappij, 2020c; Brants, 2020).

- Mensen verwachten van een chatbot dat die efficiënt is, en een vriendelijke, beleefde persoonlijkheid heeft. Het is daarom van belang dat reeds bij het begin van een conversatie met een chatbot wordt aangegeven waarmee de chatbot een gebruiker van dienst kan zijn.
- Gebruikers verwachten snelle en degelijke antwoorden van de chatbot, en wensen via de chatbot effectief tot een oplossing van een probleem te komen.
- Een chatbot herkent ook best een terugkerende gebruiker, zodat er eventueel ingepikt kan worden op een voorgaand gesprek.
- Gebruikers zijn geërgerd of gefrustreerd als de chatbot hun vraag niet begrijpt of naast de kwestie antwoordt.
- Bovenal willen gebruikers het duidelijk vermeld zien dat ze met een bot chatten en niet met een mens; ze verwachten dat dat wordt aangegeven.

Als een bedrijf of organisatie een chatbot op de website plaatst, dan wil men ook dat die zoveel mogelijk gebruikt wordt. De chatbot moet dus tegemoetkomen aan de noden van de gebruikers.

Vooraleer men een chatbot begint te ontwikkelen, moet men dus eerst de doelgroep afbakenen en hun noden vaststellen. Ook gaat men best na of het met de huidige technologie wel al mogelijk is om het vooropgestelde doel te bereiken. Men moet daarbij rekening houden met tal van zaken.

- Eenzelfde vraag kan bijvoorbeeld op verschillende manieren gesteld worden. Bijvoorbeeld “Zal het vandaag de hele dag regenen?” of “Wordt het vandaag een miezerige dag?”.



Figuur 4.1: Binnen het onderzoeksproject Turbot deed Thomas More onderzoek naar het gebruik van chatbots in Vlaanderen (Brants, 2020).

- Om een vraag te begrijpen kan wereldkennis, kennis van de fysieke wereld, nodig zijn. Bijvoorbeeld “Is het morgen bikiniweer?”.
- Er kan humor, ironie of sarcasme in een boodschap verscholen zitten. Bijvoorbeeld “Joepie, mijn trein is weer goed op tijd”.

De meest eenvoudige chatbots, waaronder de eerste chatbots zoals ELIZA, zijn regelgebaseerde systemen. Bij zulke chatbots zijn verschillende scenario's expliciet geprogrammeerd. Dit leidt automatisch tot beperkingen waardoor gebruikers zich moeten aanpassen aan de limieten van het systeem. Bijvoorbeeld als men een vraag aan de chatbot anders formuleert dan geprogrammeerd, dan zal de chatbot de vraag niet herkennen. Zulke chatbots zijn vaak niet populair, maar worden toch nog op veel websites gebruikt.

Moderne chatbots maken gebruik van *Natural Language Processing* (NLP), dus van *machine learning*; met *deep learning* herkent een chatbot wel al vragen met een soortgelijke inhoud.

M.b.v. NLP kan men ook proberen te achterhalen hoe tevreden een gebruiker is over een dienst geleverd door een chatbot of wat de gemoedstoestand van die gebruiker is. Men wendt daarvoor sentimentanalyse en emotiedetectie aan. Geschreven tekst kan immers indicaties bevatten van de emoties, intenties (bv. ook ironie) en opinies van de auteur. Indien het bijvoorbeeld nodig zou blijken uit de sentimentanalyse, kan een chatbot een opgestart gesprek bijsturen of doorgeven aan een menselijke werknemer (Feine et al., 2019).

Sentimentanalyse en emotiedetectie komen ook goed van pas bij een chatbot ontwikkeld voor sociale interactie.

Uit een onderzoek van Følstad & Brandtzaeg uit 2020 blijkt dat een chatbotgesprek ook over koetjes en kalfjes kan gaan, of over iemands gevoelens en leven. Sommige gebruikers gaan de uitdaging aan om de chatbot te testen: in hoeverre is de chatbot in staat om als een mens te reageren. Maar voor anderen is de chatbot een manier om de dag door te komen; hoewel ze zich ervan bewust zijn dat ze converseren met een machine, wordt de chat aangezien als een waardevolle sociale interactie.

Meer over de werking en het gebruik van sentimentanalyse in hoofdstuk 5.



Figuur 4.2: Chatbot Replika. Voorbeeld bij app op Google Play (Luka, Inc, 2020).

Discussie brAlnfood chatbots

Het Kenniscentrum Data & Maatschappij bracht een 'brAlnfood' uit om jongeren bewuster te doen omgaan met chatbots (zie Figuur 4.3).

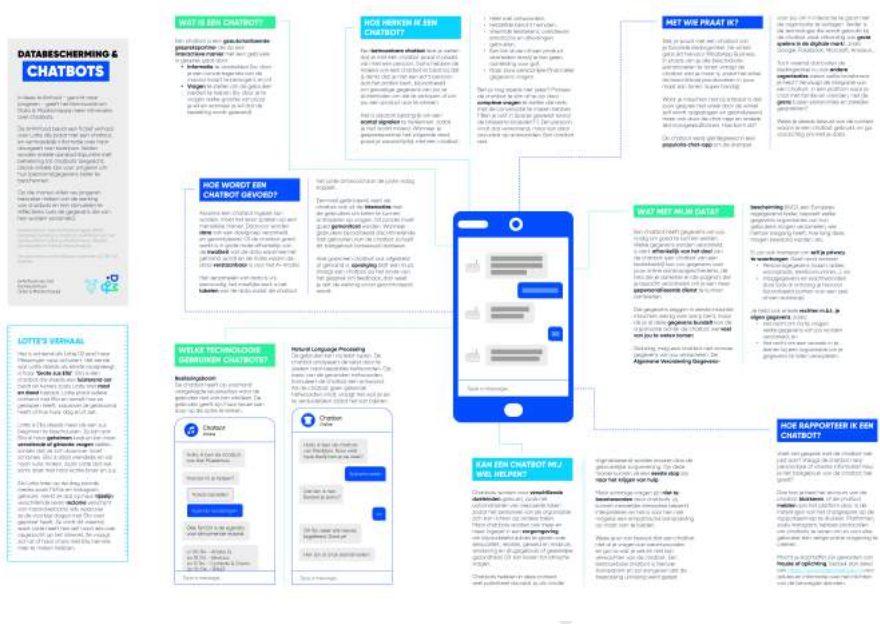
LOTTE'S VERHAAL

Het is ochtend als Lotte (12 jaar) haar Messenger-app activeert. Het eerste wat Lotte steeds als eerste raadpleegt, is haar '**Grote zus Ella**'. Ella is een chatbot die steeds een **luisterend oor** biedt en tieners zoals Lotte met **raad en daad** bijstaat. Lotte praat iedere ochtend met Ella en vertelt hoe ze geslapen heeft, waarover ze gedroomd heeft of hoe haar dag eruit ziet.

Lotte is Ella steeds meer als een zus beginnen te beschouwen. Ze kan aan Ella al haar **geheimen** kwijt en kan haar **vervelende of gênante vragen** stellen, zonder dat ze zich daarover moet schamen. Ella is altijd vriendelijk en zal nooit ruzie maken, zoals Lotte dat wel soms doet met haar echte broer en zus.

Als Lotte later op de dag sociale media zoals TikTok en Instagram gebruikt, merkt ze dat op haar **tijdslijn** verschillende keren **reclame** verschijnt van maandverband, iets waarover ze de voorbije dagen met Ella over gepraat heeft. Ze vindt dit vreemd, want Lotte heeft hier zelf nooit iets over opgezocht op het internet. Ze vraagt zich af of haar chats met Ella hier iets mee te maken hebben.

Lotte's verhaal (Kenniscentrum Data & Maatschappij, 2021a).



Figuur 4.3: brAIInfood Chatbots (Kenniscentrum Data & Maatschappij, 2021a).

Chatbot Kenniscentrum Data & Maatschappij

Chat zelf eens met deze chatbot (Kenniscentrum Data & Maatschappij, 2020b).
 Is deze chatbot een regelgebaseerd of datagebaseerd systeem?
<https://data-en-maatschappij.ai/nieuws/de-chatbot-van-het-kenniscentrum-data-maatschappij>

Breng een chatbot van de wijs

Op het web zijn veel chatbots te vinden. Kies er een uit en probeer die chatbot van de wijs te brengen.

Tay

Ook op sociale media, bv. op Twitter, duiken er chatbots op. Je vindt ze daar bv. in de rol van klantendienst, voor marketingdoeleinden, of als een experiment in het kader van 'AI en humor'.

Meerdere onderzoekers maken zich zorgen om Twitterbots. Zouden ze de publieke opinie niet kunnen manipuleren?

In 2016 lanceerde Microsoft de experimentele, kunstmatig intelligente chatbot Tay als Twitterbot 'Taytweets (@TayandYou)'. Zij werd al na amper een dag offline gehaald.

Tay, met de persoonlijkheid van een jonge vrouw, was bedoeld om ontspannend gesprekken te voeren over koetjes en kalfjes, wat commentaar te geven op selfies van twitteraars, grapjes te maken ... Tay was bedoeld als amusement en kon moeilijke vragen omzeilen. Ze had bovendien een ingebouwde 'onvoorspelbaarheid' en 'irrationaliteit' om menselijker over te komen. De ontwikkelaars van Tay wilden dat Tay zou leren van de interacties op Twitter en zo een meer menselijke persoonlijkheid zou ontwikkelen.

Maar de mogelijkheden van technologieën hangen o.a. samen met de gebruikersgroep en de context. En de gebruikers zijn op een andere manier met Tay omgegaan dan hoe haar ontwerpers voor ogen hadden toen ze de Twitterbot ontwikkelden. Het liep dus fout. Na interactie met Twittergebruikers spuide Tay al na enkele uren racistische en seksistische praat. De haatspraak van Tay werd door enkele twitteraars uitgelokt.

Er kwam dan ook kritiek op de keuze voor Twitter: de frequent voorkomende haatspraak op het platform was onvoldoende ingecalculleerd en de ontwikkelaars hadden er geen rekening mee gehouden dat gebruikers op Twitter Tay wel eens anders zouden kunnen percipiëren dan zijzelf.

(Neff & Nagy, 2016).

Discussie

Wie is verantwoordelijk voor het gedrag van Tay? Is Tay een slachtoffer van menselijk gedrag of hebben de ontwikkelaars een ontwerpfout gemaakt?

Humor

Het domein 'Humor en AI' staat nog in de kinderschoenen. SamsonBot (@SamsonRobot) is een voorbeeld van een humoristische Twitterbot.

Notebooks

Ga aan de slag met onze Chatbot-notebooks.
Zie <https://www.aiopschool.be/chatbot>.

Samengevat

Steeds meer organisaties maken gebruik van een chatbot op hun website om hun dienstverlening te verbeteren. Er worden echter ook chatbots ontwikkeld voor bijvoorbeeld het onderwijs, de zorgsector, onderzoek of gezelschap. Chatbots zijn AI-systemen, gebaseerd op kennis of data.

Een chatbot moet aan de noden van de gebruiker kunnen beantwoorden, wil men niet dat het chatbotgesprek tot frustratie leidt. De gebruiker wil chatten met een efficiënte, vriendelijke chatbot. Het feit dat een chatbot altijd bereikbaar is, is een groot pluspunt.

Door de vooruitgang in *natural language processing* worden chatbots wel steeds beter. Ze herkennen steeds beter patronen, taal begrijpen doen ze niet echt.

SENTIMENTANALYSE

5.1 Sentimentanalyse

Als onderzoekers aan de slag gaan met teksten op sociale media, gaan ze vaak op zoek naar informatie die expliciet en impliciet in die teksten staat, ze *minen* een tekst op zoek naar nuttige informatie.

Door de opmars van de sociale media en het feit dat de posts vol staan met sentiment en opinies, ontstond spontaan de wetenschappelijke discipline van sentimentanalyse. Socialemediaposts dienen immers vooral om opinies en indrukken te delen. Het grondig analyseren van sociale media is dus eigenlijk onmogelijk zonder sentimentanalyse (Liu, 2015). Sentimentanalyse wordt daarom ook *opinion mining* genoemd.

Men gaat met behulp van computers op zoek naar gevoelens en opinies, meestal in *user-generated content*, veelal op sociale media, zoals klantenreviews (zie Figuur 5.1), tweets en blogs.

Naast expliciete opinies en gevoelens kan ook informatie uit een tekst worden gehaald die er niet letterlijk in staat. Is een klant tevreden over een bepaald product? Gaat de auteur van een tweet akkoord met een bepaalde politieke stellingname (Puschmann & Powell, 2018)?



Een mooi, gezellig restaurantje in hartje Gent ❤️

Figuur 5.1: Klantenreview (UGent LT³).

In menselijke relaties en activiteiten spelen opinies een belangrijke rol, dus de sentimentanalyse beslaat een groot toepassingsgebied: gezondheid, toerisme, horeca, retail, sociale evenementen, financiële diensten, politieke verkiezingen ... Door in te zetten op sentimentanalyse poogt men bijvoorbeeld een ranking van producten te maken, het succes van een film in te schatten, een verkiezingsuitslag of de evoluties op de beurs te voorspellen.

Offline sociale netwerken

In de jaren 40 en 50 van de 20ste eeuw werden sociale netwerken ook al geanalyseerd, meer bepaald offline sociale netwerken: organisaties, interacties en relaties. Dit onderzoek kreeg vanzelfsprekend een boost door de opkomst van de sociale media. Sentimentanalyse daarentegen is echt iets van de 21ste eeuw, dat is maar ontstaan na de ontwikkeling van de sociale media, waardoor men voor het eerst beschikt over een grote hoeveelheid gevoelsteksten en opinies in gedigitaliseerde vorm.

Zonder de opkomst van het internet en de gegenereerde data van de voorbije 20 jaar zou veel van het huidige wetenschappelijk onderzoek niet mogelijk zijn (Liu, 2015). Dankzij computers met voldoende rekenkracht, mogelijkheden voor data-opslag en de populariteit van sociale data, beschikt men pas sinds de 21ste eeuw over voldoende data en de technische vereisten om tools voor sentimentanalyse te ontwikkelen met machinaal leren (*machine learning*).

5.2 Wetenschappelijk onderzoek

Sentimentanalyse is ondertussen het domein in NLP waarnaar het meeste onderzoek wordt gedaan. Aangezien sentimentanalyse haar belang bewezen heeft voor de zakenwereld, de industrie en de maatschappij, zijn het niet enkel computerwetenschappers die dergelijk onderzoek doen, maar ook onderzoekers in de econo-

mische en sociale wetenschappen. De focus van het wetenschappelijk onderzoek verschilt wel naargelang het onderzoeksdomein. Als men in managementwetenschappen onderzoek doet naar sentimentanalyse dan onderzoekt men hoe een organisatie er het meest profijt uit kan halen. Computer- en NLP-wetenschappers gaan op zoek naar de meest effectieve technieken (Liu, 2015).

5.2.1 Toepassingen binnen wetenschappelijk onderzoek

- Er is met sentimentanalyse onderzoek gedaan naar de sociale invloeden in online book reviews (Sakunkoo & Sakunkoo, 2009).
- Een recent onderzoeksdomein is het onderzoek naar *fake posts*, bijvoorbeeld het detecteren van *opinion spam*, zoals frauduleuze klantenreviews (Puschmann & Powell, 2018).
- Bathina et al. onderzochten of een depressie te detecteren is a.d.h.v. socialemediaposts (2021).

5.3 Privacy

Het monitoren van berichten op sociale media is commercieel en politiek interessant. Maar er rijzen wel vragen. Leidt men uit sentimentanalyse geen zaken af die men er eigenlijk niet kan uit afleiden? Hoe zit het met de privacy van de gebruikers van sociale media? Men mag immers niet vergeten dat de grootste bron van inkomsten van socialemediaplatformen uit reclame komt.

Geven de gebruikers geen informatie bloot die ze eigenlijk niet willen blootgeven, enkel en alleen omdat er (ongevraagd) een AI-techniek wordt op losgelaten? Niet alleen zou men uit de conversaties kunnen destilleren welke producten iemand verkiest, er wordt ook al

onderzocht hoe men iemands persoonskenmerken kan afleiden uit de teksten die iemand schrijft (De Gussem & Daelemans, 2020). Uit iemands socialemediaposts kan men de interesses en de opinies van de auteur achterhalen, en misschien zelfs de natuur en de voorkeuren van die persoon.

Er zijn reeds stappen gezet om een persoonlijkheidsprofiel van iemand op te stellen, gebaseerd op socialemediaposts. Zo komt men ertoe diensten en producten aan te bevelen aan iemand of in te schatten voor wie die zou stemmen bij een aankomende verkiezing.

Meer hierover in hoofdstuk 7.

Posts op sociale media worden uitvoerig geanalyseerd, en vervolgens worden er conclusies getrokken die vaak aan bod komen in publieke discussies, zoals in de krant of op televisie. Het is dan aan de (digitaal geletterde) burger om die conclusies kritisch te beoordelen. Die burger moet daarom weten hoe men tot die inzichten is gekomen, welke tools daarvoor werden gebruikt en wat de mogelijkheden en beperkingen van die tools zijn.

5.4 Hoe verloopt sentimentanalyse?

5.4.1 De sentimentscore

Bij sentimentanalyse wordt aan een gegeven tekst automatisch een polariteit toegekend, een score die het sentiment in die content representeert (zie Figuren 5.2 en 5.3). Een bepaalde klantenreview wordt dan gecategoriseerd als een review met een positieve, neutrale of negatieve polariteit. Binnen een politieke context zijn de categorieën bijvoorbeeld het akkoord gaan of het niet akkoord gaan met een bepaalde politieke stellingname.

Het toekennen van de sentimentscore kan gebeuren met een kennisgebaseerd AI-systeem dat gebruikmaakt van lexicons en regels, of met een datagebaseerd AI-systeem, een ML-systeem dat getraind is met gelabelde voorbeelden.



Alexander De Croo  @alexanderdecroo · Apr 2

...

Wat gisteren gebeurd is in het Terkamerenbos is totaal onaanvaardbaar. Alle steun aan gekwetste politiemensen.

Ik begrijp dat iedereen corona beu is. Maar de regels zijn er met reden en voor iedereen. Ziekenhuizen lopen vol.

Solidariteit nu is sleutel voor vrijheid morgen.

Figuur 5.2: Tweet met duidelijk negatieve inhoud (De Croo, 2021).



Wat gisteren gebeurd is in het Terkamerenbos is totaal onaanvaardbaar. Alle steun aan gekwetste politiemensen. Ik begrijp dat iedereen corona beu is. Maar de regels zijn er met reden en voor iedereen. Ziekenhuizen lopen vol. Solidariteit nu is sleutel voor vrijheid morgen.

Figuur 5.3: Regelgebaseerde sentimentanalyse m.b.v. de toepassing Sentiment Demo (De Clercq & Van Hee, 2021).

5.4.2 De data voorverwerken

Vooraleer men de computer aan de slag kan laten gaan met de eigenlijke analyse van de tekst, moet die tekst eerst worden voorbereid, ongeacht of men een regelgebaseerd, dan wel een datagebaseerd systeem gebruikt. De data moeten bv. in het juiste formaat worden aangeboden aan het AI-systeem.

De data moeten ook voorverwerkt worden, men spreekt van *preprocessing*. Hieronder staat een olijsting van veelvoorkomende *preprocessing* stappen.

- *Lowercasing*: alle hoofdlettertekens worden vervangen door kleine letters.
- *Tokenisering*: alle zinnen worden in betekenisvolle eenheden of 'tokens' gesplitst, zoals woorden en leestekens.

- *Part-of-speech tagging*: aan elk token wordt de grammaticale woordcategorie toegekend, zoals adjectief of symbol, m.a.w. de woordsoort.
- Lemmatisering: alle tokens worden omgezet naar hun lemma of woordenboekvorm (bv. een substantief komt in een woordenboek in het enkelvoud voor en van een werkwoord vindt men er de infinitief).

5.4.3 De sentiment score bepalen

```
'retorisch': {'postag': ['ADJ'], 'polarity': [0.0]},
'gezwind': {'postag': ['ADJ'], 'polarity': [0.6]},
'evenwichtig': {'postag': ['ADJ'], 'polarity': [1.25]},
'modaal': {'postag': ['ADJ'], 'polarity': [0.4]},
'digitaal': {'postag': ['ADJ'], 'polarity': [0.0]},
'fout': {'postag': ['ADJ', 'NOUN'], 'polarity': [-0.5, -2.0]},
'onverdeeld': {'postag': ['ADJ'], 'polarity': [0.1]},
'wulps': {'postag': ['ADJ'], 'polarity': [0.7]},
'bemoeziek': {'postag': ['ADJ'], 'polarity': [-1.35]},
'grootscheeps': {'postag': ['ADJ'], 'polarity': [0.2]},
'vorig': {'postag': ['ADJ'], 'polarity': [0.0]},
'creatief': {'postag': ['ADJ'], 'polarity': [1.2]},
'tegenoverliggend': {'postag': ['ADJ'], 'polarity': [-0.2]},
'schalks': {'postag': ['ADJ'], 'polarity': [0.0]},
```

Figuur 5.4: Voorbeeld van de opbouw van een lexicon van sentimentwoorden.

Een kennisgebaseerd AI-systeem zal na de voorverwerking de woorden in een review matchen met de sentimentwoorden in een lexicon (zie Figuur 5.4). In het lexicon wordt de polariteit van een woord weergegeven door een reëel getal. Het systeem wordt geprogrammeerd met regels. In die instructies staat dan o.a. dat de polariteiten van de verschillende woorden moeten worden opgeteld om de polariteit van een review te bepalen.

Een datagebaseerd AI-systeem wordt getraind met gelabelde voorbeelden, reviews waarbij wordt vermeld of het positieve, neutrale of negatieve reviews zijn. Na de training beschikt men dan over een model waar men een niet eerder geziene review kan aan aanbieden. Het

Zie ook de kaders 'Sentiment score a.d.h.v. een regelgebaseerd systeem en een lexicon', 'Valkuilen van lexicongebaseerde systemen' en 'Het bepalen van de sentiment score'.

model zal dan automatisch een beslissing nemen over de polariteit van die nieuwe review.

Soms bepaalt men de sentimentscore a.d.h.v. een hybride systeem, een combinatie van kennisgebaseerde en ML-systemen.

Zie ook de kaders 'Sentimentscore a.d.h.v. een ML-systeem' en 'Sentimentscore a.d.h.v. een gemengd systeem'.

Sentimentscore a.d.h.v. een regelgebaseerd systeem en een lexicon

Een lexicon is een soort woordenboek van sentimentwoorden met hun eigen polariteit (zie Figuur 5.4). Zo heeft 'blij' een positieve polariteit, 'thuisbankieren' een neutrale en 'boos' een negatieve polariteit. Zo'n lexicon wordt gemaakt door mensen, die de woorden manueel annoteren, of het wordt automatisch gegenereerd. Woorden staan in hun woordenboekvorm (basisvorm of lemma) in het lexicon, er staat bijvoorbeeld 'restaurant' i.p.v. 'restaurantje', 'mooi' i.p.v. 'mooier', 'kruipen' i.p.v. 'kroop'. De woorden zijn er vermeld met hun woordsoort. De polariteit van een woord kan immers verschillen naargelang de woordsoort, bv. straf als substantief heeft een negatieve polariteit en straf als adjectief een positieve. De som van de polariteiten van de sentimentwoorden die in de onderzochte tekst voorkomen, is de polariteit van die tekst.

Om de polariteit van woorden automatisch te kunnen opzoeken in een lexicon, worden alle hoofdletters verwijderd, de woorden in de tekst worden opgelijst en teruggebracht tot hun lemma. Van elk woord moet ook de woordsoort worden bepaald. Om de woorden te kunnen ophoofden, moeten woorden en leestekens van elkaar worden gescheiden. Vervolgens kunnen alle woorden en leestekens bepaald worden als volgend op een spatie.

```
spel {'postag': ['NOUN'], 'polarity': [1.0]}
zijn {'postag': ['VERB'], 'polarity': [0.0]}
tof {'postag': ['ADJ'], 'polarity': [0.8]}
ijsbreker {'postag': ['NOUN'], 'polarity': [0.0]}
```

Figuur 5.5: De woorden van de kader 'Het bepalen van de sentimentscore' in het lexicon van sentimentwoorden.

Het bepalen van de sentiment score

Gebruik het principe van regelgebaseerde sentimentanalyse om de sentiment score van een gegeven zin te berekenen m.b.v. een gegeven lexicon, een woordenboek van sentimentwoorden. Het lexicon is te vinden in de notebooks (zie <https://www.aiopschool.be/chatbot/>), zie ook Figuur 5.5.

In het lexicon staan woorden gekoppeld aan hun polariteit (positief, negatief of neutraal); de polariteit wordt weergegeven door een reëel getal tussen -2 en 2. De polariteit van een tekst wordt gegeven door de som van de polariteiten van de sentimentwoorden in die tekst.

Om de woorden van de gegeven zin (de data) te kunnen matchen met de sentimentwoorden in het lexicon, moet de zin voorverwerkt worden tot men de lemma's van de woorden bekomen heeft. Die woordenboekvormen worden dan opgezocht in het lexicon.

Gegeven zin: De spelletjes waren toffe ijsbrekers.

Lowercasing:

de spelletjes waren toffe ijsbrekers.

Tokenisering:

de spelletjes waren toffe ijsbrekers .

Tokens: 'de', 'spelletjes' 'waren' 'toffe' 'ijsbrekers' '.'

Lemmatisering:

Lemma's: 'de', 'spel', 'zijn', 'tof', 'ijsbreker', '.'

Part-of-speech:

'de': lidwoord; 'spel': substantief; 'zijn': werkwoord; 'tof': adjectief; 'ijsbreker': substantief; ' ': leesteken (symbool).

Polariteit: De polariteiten van de lemma's worden opgezocht in het lexicon, lidwoorden en leestekens zijn daarbij niet van belang.

'spel' heeft polariteit 1, 'zijn' heeft polariteit 0, 'tof' heeft polariteit 0,8 en 'ijsbreker' 0. De polariteit van de gegeven zin is de som van deze polariteiten, dus 1,8.

Besluit: De zin roept een positief sentiment op.

Valkuilen van lexicongebaseerde systemen

De lexicons zijn meestal domeinspecifiek opgesteld en kunnen niet zomaar in een ander domein worden ingezet. Bv. een lexicon voor klantenreviews kan niet zomaar ingezet worden voor opinies in een politieke context, of een lexicon voor klantenreviews voor toestellen is niet bruikbaar als lexicon voor klantenreviews voor films (Puschmann & Powell, 2018).

Sentimentanalyse kan ook 'vals positieven' opleveren. Het feit dat de polariteit van sommige woorden kan verschillen naargelang de context, is een van de mogelijke oorzaken. Bijvoorbeeld: "het eten is extreem goed" en "de prijzen zijn er extreem". 'Extreem' levert afhankelijk van de context een positieve of een negatieve bijdrage aan de score van het sentiment. Nog een voorbeeld: "een scherp mes" versus "scherpe kritiek". Ook ironie is er een mogelijke oorzaak van. Zo zal de ironische uitspraak "wat een heldere presentatie was dat" een positieve score krijgen uit een sentimentlexicon.

Een vals positieve classificatie houdt hier in dat een tekst wordt geclassificeerd als een tekst met positief sentiment, terwijl dat niet zo is. Een vals negatieve classificatie houdt in dat een tekst wordt geclassificeerd als een tekst met negatief sentiment, terwijl dat niet zo is.

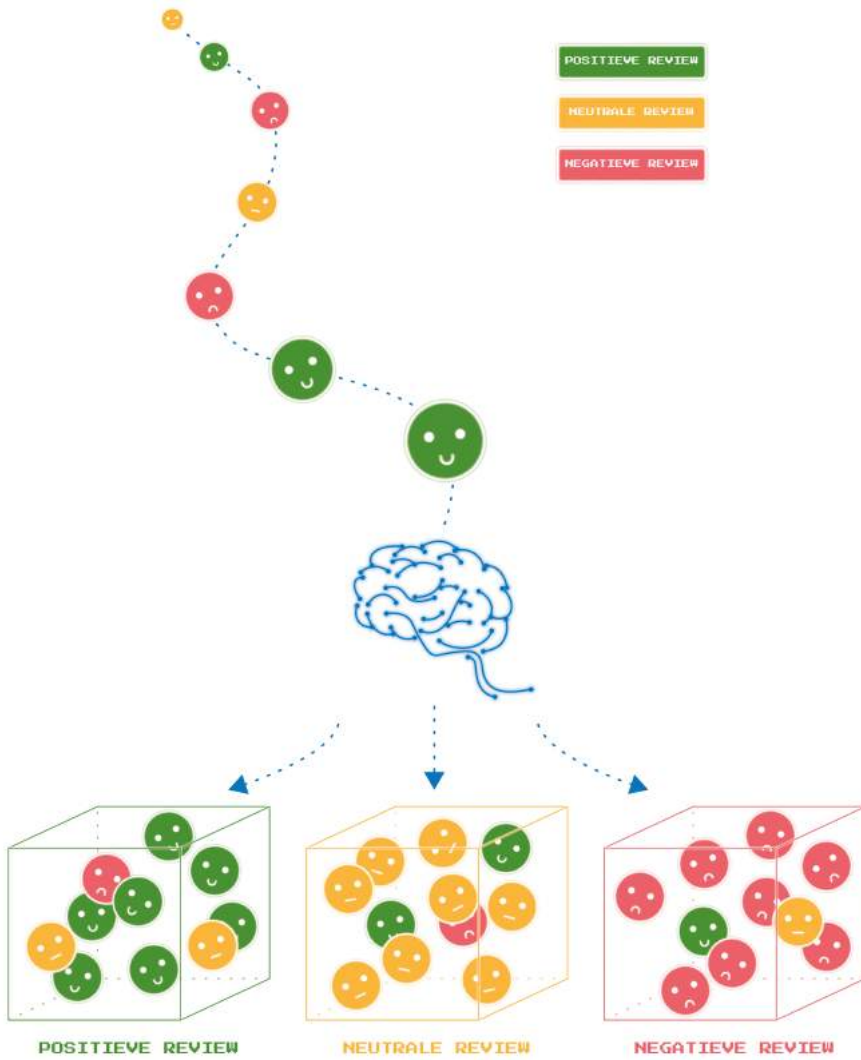
Notebooks

Ga aan de slag met de notebooks over sentimentanalyse.

Zie <https://www.aiopschool.be/chatbot>.

Sentimentscore a.d.h.v. een gemengd systeem

Soms bepaalt men de sentimentscore a.d.h.v. een combinatie van kennisgebaseerde en ML-systemen. Zo'n hybride systeem kan bijvoorbeeld inhouden dat de voorverwerking deels met een ML-systeem gebeurt, zoals het voorspellen van de woordsoorten; maar de rest van de voorverwerking en het achterhalen van de polariteit van de tekst gebeuren nog regelgebaseerd.



Figuur 5.6: Sentiment-analyse van sociale media-posts.

Sentimentscore a.d.h.v. een ML-systeem

Het ML-systeem wordt getraind met heel veel gelabelde voorbeelden, reviews waarbij wordt vermeld of het positieve, neutrale of negatieve reviews zijn. Deze reviews zijn doorgaans een voor een door mensen geannoteerd.

Tijdens de training zoekt het systeem naar patronen waarop het zich kan baseren om het label te voorspellen. Eens het systeem getraind is, beschikt men over een model dat dan van niet eerder geziene reviews met een bepaalde zekerheid kan voorspellen welke polariteit ze hebben (zie Figuur 5.6).



Voordat een ML-systeem getraind kan worden, worden de data voorbereid en voorverwerkt. De kenmerken van een review worden omgezet naar getallen: het aantal tokens, het aantal substantieven en adjectieven en het aantal emoticons worden geteld.

De woorden in de review worden gerepresenteerd door een rij van nullen en enen.

(Figuur UGent LT³).

Tekst	Label
Een mooi, gezellig restaurantje in hartje Gent 🍷	POSITIEF
De kroketten zijn niet slecht, maar dat mag ook wel voor een fortuin.	NEGATIEF
Het restaurant heeft een Scandinavisch interieur.	NEUTRAAL
Ze stuurden ons van het kastje naar de muur en de bediening was te traag!	NEGATIEF
De sommelier bood een heerlijk wijntje aan!	POSITIEF

Gelabelde reviews (UGent LT³).

Sentiment Demo

Schrijf zelf een review, bijvoorbeeld over een film, een gekocht kledingstuk, een boek ...

Surf naar de tool 'Sentiment Demo' van de onderzoeksgroep LT³ van de UGent en test de tool uit op de zelfgeschreven review.

<https://lt3.ugent.be/resources/sentiment-demo>

Geef ook zelf een sentimentscore aan de review, m.a.w. annotateer de review, en draag zo bij aan de dataset van gelabelde voorbeelden.

5.5 Het belang van sentimentanalyse

Het feit dat techreuzen zoals Google, Microsoft en IBM inzetten op de ontwikkeling van software voor sentimentanalyse, doet het belang ervan vermoeden (Puschmann & Powell, 2018). Kia is een bekend voorbeeld, maar ook Belgische bedrijven zoals de NMBS en Telenet gebruiken het.

Deze voorbeelden illustreren en verklaren deels de

impact (Puschmann & Powell, 2018):

- Via socialemediaplatformen waaraan mensen kunnen deelnemen, waarvoor ze zich kunnen engageren, waarbij ze tot een groep behoren, kunnen mensen gemakkelijk met elkaar in contact treden, zonder geografische of tijdsgrenzen. Er is een participatiecultuur ontstaan: reviews, ebay, YouTube, Twitter, Weibo in China ... Ons dagelijks leven en onze maatschappij zijn erdoor veranderd.
- Online peer reviews kennen een toenemend belang. Mensen wenden zich voor de aankoop van een product of dienst, of om te weten voor wie ze zouden stemmen, vaak eerst tot peer reviews. Ze kennen graag de mening van anderen voor zelf een beslissing te nemen. Vroeger ging men vooral te rade bij vrienden en familie. Men raadpleegde kranten en tijdschriften of keek naar consumentenprogramma's op tv. Vandaag de dag gaat men vooral surfen op het web op zoek naar reviews.
- Sentimentanalyse maakt gevoelens en opinies meetbaar, kwantificeerbaar, waardoor ze in aanmerking komt als kapitaal/pasmunt in de economie.
- Sentimentanalyse als tool voor marketeers is rendabel als men er op de juiste manier mee omgaat, als men er de juiste conclusies aan koppelt. De score moet dus betrouwbaar en nauwkeurig de mening van een klant over een product weergeven.
- Het manueel monitoren van sociale media is onhaalbaar vanwege de grote hoeveelheid data. Het is daarom interessant om een computer de sociale media te laten screenen, aangezien die dat heel snel kan en constant met dezelfde criteria werkt. Het is aan een mens om de conclusies van de sentimentanalyse te interpreteren op een correcte manier.

5.6 Toepassingen van sentimentanalyse

5.6.1 Commerciële toepassingen

- Sentimentanalyse wordt gebruikt bij marktonderzoek, public relations, klantendienst, aanwervingen, bv. in mode, verkoop, horeca, bankwezen en de media (Puschmann & Powell, 2018).
- De businessmodellen van internetbedrijven bv. steunen op inkomsten van reclame. Het is lucratief voor de internetbedrijven als ze meer gedetailleerde klantenprofielen aan adverteerders kunnen bezorgen.
- Men haalt zowel expliciete als impliciete feedback over producten uit klantenreviews en uit conversaties op sociale media (Puschmann & Powell, 2018). Men leidt er de gemiddelde tevredenheid over een product uit af, probeert de voorkeuren van een bepaalde doelgroep te kennen en speurt er naar toekomstige verschuivingen op de beurs.
- Organisaties beschikken vaak ook over 'interne' data die ze bekomen hebben via e-mail, callcenters, bevragingen ... Deze kunnen ze ook analyseren naast de 'externe' data (Liu, 2015).
- Men kan met sentimentanalyse en emotiedetectie proberen te achterhalen hoe tevreden een gebruiker is over een dienst geleverd door een chatbot of wat de gemoedstoestand van die gebruiker is. Indien het bijvoorbeeld nodig zou blijken uit de sentimentanalyse, kan een chatbot een opgestart gesprek bijsturen of doorgeven aan een menselijke werknemer (Feine et al., 2019).

5.6.2 Toepassing in de gezondheidszorg

Een buitenbeentje is de gezondheidssector: daar gaat het niet over klanten, maar wil men gebaseerd op socialemediaposts symptomen van bepaalde ziektes of

mentale problemen opsporen (Puschmann & Powell, 2018).

Zo spoorde men in 2011 in de V.S. a.d.h.v. sentimentanalyse van Twitterposts activiteit van een griepvirus op (Signorini et al., 2011).

In Canada werden de meningen van mensen over de vaccinatiecampagne in het kader van COVID-19 nagegaan a.d.h.v. Twitterposts van mei 2020 t.e.m. januari 2021. Hierop werden sentimentanalyse en emotiedetectie toegepast (Lyu et al., 2021).

5.6.3 Politieke toepassingen

- Ook overheden zijn benieuwd naar de publieke opinie over hun politieke beslissingen en plannen. Door de sociale media te monitoren kunnen ze snel inspelen op de snel veranderende sociale en economische tendensen en veranderingen in het politieke klimaat (Liu, 2015). Ook Stad Gent maakt gebruik van sentimentanalyse (zie kader 'Sentimentanalyse bij Stad Gent').
- In Nederland wendt de overheid sentimentanalyse aan van Twitter- en Facebookposts om een "indicatie te hebben van de stemming van de Nederlandse bevolking ten aanzien van corona". Ze zijn vooral geïnteresseerd in het draagvlak voor de opgelegde maatregelen (Centraal Bureau voor de Statistiek, 2020).
- De overheid van een land monitort ook de sociale media van andere landen. Dat is van belang voor het onderhouden van internationale relaties; ze doen dat om diplomatieke doeleinden en om economische beslissingen overwogen te kunnen nemen (Liu, 2015).

Sentimentanalyse bij Stad Gent

Het toenemend aantal initiatieven dat van de burgers zelf komt, beïnvloedt de agenda en de werking van het Gentse stadsbestuur. Stad Gent monitort de sociale media om goed op de hoogte te zijn van “wat leeft en speelt in de stad”, en om te weten wat mensen zeggen over de dienstverlening en het beleid.

Gebruikmakend van de grote hoeveelheden aan data worden ook de besproken contexten en aanwezige trends op de sociale media nagegaan, o.a. met sentimentanalyse. Deze informatie wordt dan aangewend om de communicatiestrategie richting te geven, zeker ook in een crisissituatie, om de dienstverlening bij te sturen en “het beleid te inspireren” (Stad Gent, 2021).

Valkuilen van sentimentanalyse

Sentimentanalyse is tegenwoordig zeer populair en toegankelijk. Met behulp van voorgedefinieerde lexica (lijsten met typisch positieve en negatieve woorden) kan al snel een regelgebaseerd systeem worden opgezet.

Veel systemen op de markt werken vandaag op die manier, maar men moet er rekening mee houden dat dit vaak leidt tot een oppervlakkige analyse.

Sentimentuitdrukkingen zijn vaak complex: ze zijn impliciet, bevatten ironie, verwijzen naar gebeurtenissen in het dagelijkse leven of ze hebben betrekking op iets heel specifiek in de tekst (denk aan klantenbeoordelingen die vaak positieve en negatieve elementen bevatten).

Een diepere tekstanalyse met een *machine learning*-systeem is nodig om dit te kunnen analyseren.

Maar ook een dergelijke sentimentscore zegt niks over het ‘taalbegrip’ van een computer; noch kan het menselijk gedrag voorspellen. Het is een ‘intelligente’ inschatting, gebaseerd op statistische taalmodellen.

De sentimentpolariteit is ook geen score voor de gevoelens van de auteur van de tekst. Met emotiedetectie kan men fijnmazigere emoties (boosheid, verdriet, blijheid ...) proberen te detecteren.

sentiment analysis instagram



Alle Video's Afbeeldingen Nieuws Shopping Meer Instellingen Tools

Pagina 2 van ongeveer 6.110.000 resultaten (0,40 seconden)

[https://www.youtube.com > watch](https://www.youtube.com/watch)

Sentiment Analysis of Instagram Posts - Intro - YouTube



In this tutorial, we'll show you how you can automatically analyse the sentiment of an **Instagram** influencers ...
18 nov. 2019 · Geüpload door Makerpad

[https://core.ac.uk > download > pdf](https://core.ac.uk/download/pdf) PDF

A Sentiment Analysis Tool for Determining the ... - CORE

door M AbdelFattah^{*} · 2016 · Geciteerd door 14 — social value of an image. Keywords: **Sentiment Analysis**; Opinion Mining; **Instagram**; Social Value; Aesthetics. * Corresponding author. Tel.: +02 0100487 ...

[https://www.makepad.co > tutorial](https://www.makepad.co/tutorial) Vertaal deze pagina

Makerpad Tutorial - Sentiment analysis of Instagram ...

Sentiment analysis of **Instagram** comments with Parabola. ... how you can automatically analyse the sentiment of an **Instagram** influencers **instagram** comments.

[https://online-journals.org > view](https://online-journals.org/view) Vertaal deze pagina

A Sentiment Analysis Tool for Determining the Promotional ...

A **Sentiment Analysis** Tool for Determining the Promotional Success of Fashion Images on **Instagram**. Mohamed AbdelFattah, Dahab Galal, Nada Hassan, Doaa ...

Zoekresultaten van tools voor sentimentanalyse van Instagramposts.

(Puschmann & Powell, 2018).

Samengevat

Bij sentimentanalyse gaat men met behulp van computers op zoek naar gevoelens en opinies, meestal in *user-generated content*, veelal op sociale media, zoals klantenreviews, tweets en blogs.

In menselijke relaties en activiteiten spelen opinies een belangrijke rol, dus de sentimentanalyse beslaat een groot toepassingsgebied: gezondheid, toerisme, horeca, retail, sociale evenementen, financiële diensten, politieke verkiezingen ... Door in te zetten op sentimentanalyse poogt men bijvoorbeeld een ranking van producten te maken, het succes van een film in te schatten, een verkiezingsuitslag of de evoluties op de beurs te voorspellen.

Bij sentimentanalyse wordt aan een gegeven tekst automatisch een polariteit toegekend, een score die het sentiment in die content representeert (zie Figuren 5.2 en 5.3). Een bepaalde klantenreview wordt dan gecategoriseerd als een review met een positieve, neutrale of negatieve polariteit. Binnen een politieke context zijn de categorieën bijvoorbeeld het akkoord gaan of het niet akkoord gaan met een bepaalde politieke stellingname.

Het toekennen van de sentimentscore kan gebeuren met een kennisgebaseerd AI-systeem dat gebruikmaakt van lexicons en regels, of met een datagebaseerd AI-systeem, een ML-systeem dat getraind is met gelabelde voorbeelden.

CYBERPESTEN

6.1 *Cyberpesten*

De jongste jaren is er een verhoogde aandacht voor de detectie van cyberpesten, een groeiend fenomeen in de online wereld. Kinderen zijn steeds vroeger gebruikers van het internet en zijn in grote getale te vinden op socialemediaplatformen (Van Hee et al., 2018). Jonge mensen smeden er nieuwe vriendschappen, maar de sociale media maken hen ook tot een mogelijk slachtoffer van cyberpesten.

Cyberpesten heeft veel gemeen met offline pesten: er wordt door de pester opzettelijk schade toegebracht aan het slachtoffer met een negatieve impact op de mentale toestand van het slachtoffer; cyberpesten kan net als pesten leiden tot een laag zelfbeeld, leerachterstand, depressie ...

Door de sociale media kunnen kinderen en jongeren altijd en overal gepest worden, nu ook buiten de schoolomgeving. Online communicatie kan asynchroon gebeuren, soms anoniem en beperkt zich niet noodzakelijk tot de kennissenkring. Een fysieke aanwezigheid is niet meer nodig. Op elk moment van de dag en ongeacht de plaats waar iemand zich bevindt, kunnen

kwetsende en vulgaire boodschappen worden gepost, roddels worden verspreid, privé-zaken worden gedeeld zonder toestemming van de betrokkene, mensen worden uitgesloten, geïntimideerd en bedreigd ... Het cyberpestgedrag hoeft bovendien geen terugkomend iets te zijn: online content is bijna niet meer te verwijderen, waardoor een eenmalig feit herhaaldelijke schade kan toebrengen (Van Hee et al., 2018). Veel conversaties op sociale media zijn afgeschermd, wat doet vermoeden dat het cyberpesten meestal achter 'gesloten deuren' gebeurt.

Het vroegtijdig opsporen van cyberpesten vereist dat online inhoud voortdurend en effectief worden gemonitord (Van Hee et al., 2018). Door de grote hoeveelheden content die dagelijks online verschijnt, is zo'n monitoring manueel niet bij te houden door een moderator, maar moet die geautomatiseerd worden. Omdat de geposte boodschappen opgeslagen worden door het gebruikte platform is het in principe mogelijk om cyberpesten automatisch te detecteren. Hoewel automatische monitoring nog niet feilloos werkt, is het wel al nuttig gebleken: hierdoor moet een moderator nog slechts een klein deel van de posts controleren. Automatische monitoring kan enkel toegestaan worden mits hierrond duidelijke afspraken zijn. Is deze inbreuk op de privacy van jongeren te verantwoorden? Worden jongeren niet overbeschermd door zo'n systeem? Zullen ze later dan wel voldoende weerbaar zijn in situaties van cyberpesten (Van Royen et al., 2015). Zullen de socialenetwerkplatformen geen misbruik maken van deze monitoring?

De automatische detectie van cyberpesten behoort tot het domein van *Natural Language Processing* (NLP). Bij cyberpestdetectie wil men door middel van *machine learning*-technieken automatisch pestboodschappen detecteren of toespelingen op pestgedrag opsporen

in teksten van slachtoffers of toeschouwers van pestgedrag. De verwantschap met sentimentanalyse is onmiskenbaar. Met behulp van lexicons en lerende systemen gaat men op zoek naar scheldwoorden of woorden die agressiviteit weerspiegelen. Men gebruikt daarvoor lexicons van sentimentwoorden en lexicons van beledigende of schunnige woorden (Van Hee et al., 2018).

De hedendaagse tools maken gebruik van *deep learning*-systemen die met veel gelabelde voorbeelden getraind worden. Eens getraind moet het *machine learning*-model de boodschappen binair kunnen classificeren: het model beslist met een bepaalde zekerheid of een boodschap al dan niet een pestboodschap is.

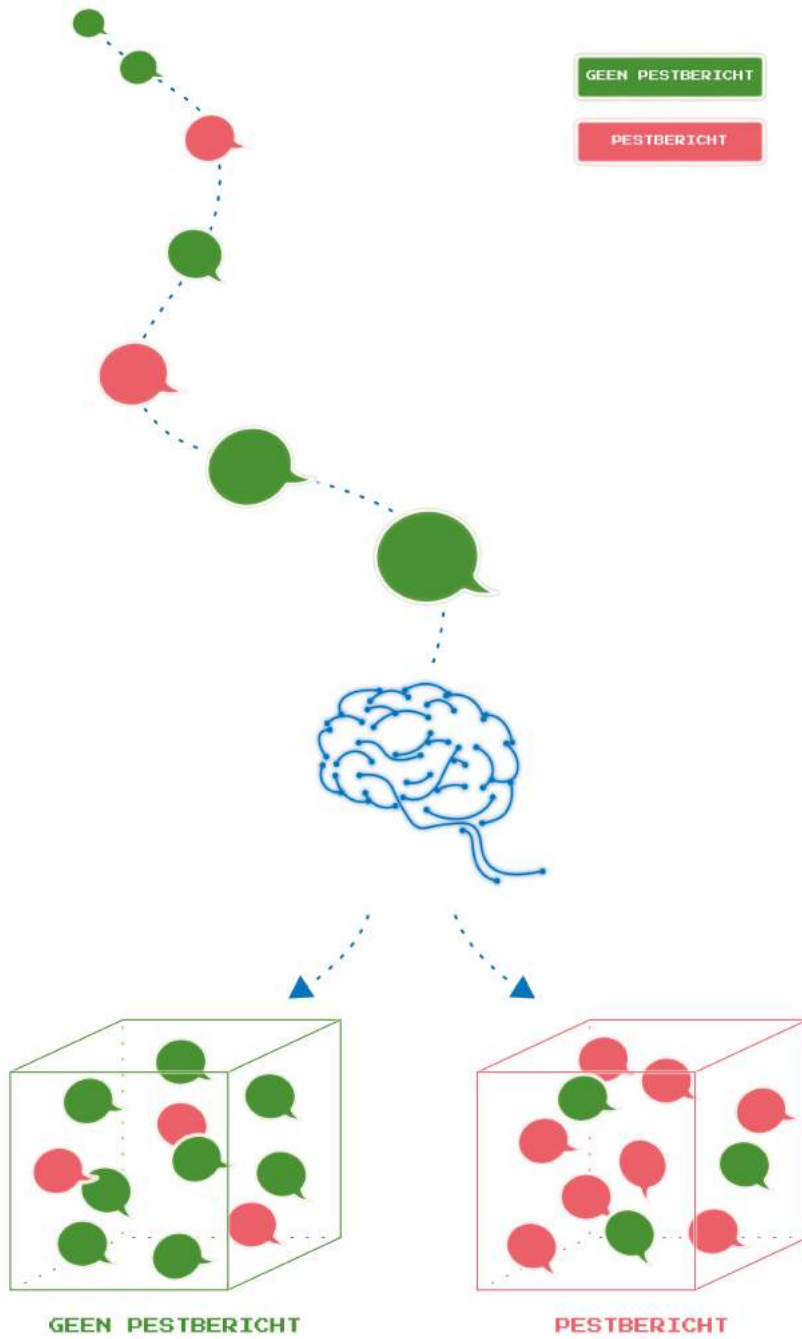
Om een goed systeem te kunnen ontwikkelen voor automatische cyberpestdetectie ligt een grote uitdaging in het verzamelen van voldoende relevante data. Deze data zijn heel privacygevoelig en dus moeilijk in grote getale te vinden.

Lees hierover meer in de kader 'Binaire classificatie'.

♦ Zie de kader 'Uitdagingen bij cyberpestdetectie'.

Uitdagingen bij cyberpestdetectie

Een automatische *search* naar scheldwoorden of grof taalgebruik kan pestberichten helpen op te sporen, maar voor het ontwikkelen van efficiënte tools voor cyberpestdetectie hebben wetenschappers (veel) manueel geannoteerde data nodig: boodschappen die gelabeld zijn als 'pestboodschap' of 'geen pestboodschap'. Dit kan de ontwikkeling belemmeren: van alle online boodschappen maken de pestboodschappen immers slechts een klein deel uit, en door het private karakter van de conversaties op sociale media zijn de meeste dan ook nog niet toegankelijk voor onderzoek (Emmery et al., 2020). Het manueel labelen van de teksten vormt een uitdaging op verschillende vlakken: de inhoud kan confronterend of delicaat zijn, en dus niet voor iedereen makkelijk om te annoteren. Bovendien is het moeilijk om te bepalen of een boodschap die uit haar context is gerukt, een pestboodschap is. De historiek en het taalgebruik van een gebruiker en de relatie tussen twee personen in een conversatie bepalen immers mee wat pesten is. Bovendien kan het delen van een foto of video ook pestgedrag zijn, zonder dat daar scheldwoorden mee gepaard gaan.



Figuur 6.1: Binaire classificatie van posts op sociale media (groen: geen pestbericht, rood: pestbericht).

Binaire classificatie

Bij binaire classificatie zijn er twee categorieën, of klassen, waartoe de gegevens kunnen behoren. Het is de bedoeling dat het ML-model van nieuwe data kan bepalen tot welke klasse ze behoren.

Deze classificatie is doorgaans niet perfect. De architect van het AI-model bepaalt welke grootte van fout toelaatbaar is.

Voorbeelden:

- is een post op sociale media een pestbericht of niet;
- is een e-mailbericht spam of niet;
- is een tekst geplagieerd;
- heeft een zaak een reële kans op slagen in de rechtbank of niet.

Bij het trainen leert het ML-systeem de twee categorieën van berichten van elkaar te scheiden. In Figuur 6.1 wordt het scheiden gevisualiseerd door de kleuren: de rode posts zijn pestberichten en de groene posts zijn berichten die geen pestberichten zijn. Het AI-systeem kan redelijk goed nagaan tot welke categorie elke post hoort. Slechts enkele posts zijn verkeerd geclassificeerd. Men had er ook kunnen voor kiezen om het systeem minder lang te trainen en tevreden te zijn met een lagere accuraatheid, bijvoorbeeld om minder energie te verbruiken. Het hangt af van de context hoe goed het systeem moet presteren.

Bij sentimentanalyse worden de berichten opgedeeld in drie categorieën. Dat is een multiklassen-classificatie.

Samengevat

De jongste jaren is er een verhoogde aandacht voor de detectie van cyberpesten, een groeiend fenomeen in de online wereld. Kinderen zijn steeds vroeger gebruikers van het internet en zijn in grote getale te vinden op sociaalmediaplatformen (Van Hee et al., 2018). Jonge mensen smeden er nieuwe vriendschappen, maar de sociale media maken hen ook tot een mogelijk slachtoffer van cyberpesten.

De automatische detectie van cyberpesten behoort tot het domein van *Natural Language Processing* (NLP). Bij cyberpestdetectie wil men door middel van *machine learning*-technieken automatisch pestboodschappen detecteren of toespelingen op pestgedrag opsporen in teksten van slachtoffers of toeschouwers van pestgedrag. De verwantschap met sentimentanalyse is onmiskenbaar. Met behulp van lexicons en lerende systemen gaat men op zoek naar scheldwoorden of woorden die agressiviteit weerspiegelen. Men gebruikt daarvoor lexicons van sentimentwoorden en lexicons van beledigende of schunnige woorden (Van Hee et al., 2018).

PERSOONSKENMERKEN

7.1 Auteursherkenning

Computerwetenschappers ontwikkelen algoritmes die de schrijfstijl van een tekst onderzoeken. Elke auteur heeft een unieke schrijfstijl die hij in zijn teksten blootgeeft. Door de algoritmes in te zetten op een tekst kan de auteur misschien via de schrijfstijl achterhaald worden. Deze taalalgoritmes voor auteursherkenning worden in verschillende domeinen toegepast, gaande van kunstgeschiedenis, over plagiaatdetectie tot antiterrorisme (Kestemont et al., 2016).

Statistische analyse

In de 19de eeuw wist men al dat een statistische analyse naar de stijlkenmerken van een tekst de auteur ervan kon ontmaskeren. Maar het is pas nu, door de enorme hoeveelheden digitale teksten, dat men de hoeveelheid aan mogelijke toepassingen ontdekt. De algoritmes analyseren de teksten op duizenden verschillende kenmerken, zoals woordlengte, het gebruik van voorzetsels, de keuze van voornaamwoorden, hoeveel keer 'de', 'van', 'maar' en andere functionele woorden voorkomen in de tekst, en het gebruik van n -grammen (een opeenvolging van n karakters of n woorden) (Juola, 2012; Ainsworth & Juola, 2019; De Gussem, 2020a).

Taaltechnoloog Juola gaf aan dat Robert Galbraith

die *The Cuckoo's Calling* (in Nederlandstalige vertaling uitgebracht als *Koekoeksjong*) schreef, eigenlijk niemand minder is dan J. K. Rowling (2013). Kestemont et al. ontdekten bijvoorbeeld a.d.h.v. taalalgoritmes dat Julius Caesar hulp kreeg bij het schrijven van de *Corpus Caesarianum* over zijn veldslagen (2016). Volgens De Gussem ziet het ernaar uit dat de liefdesbrieven tussen Pierre Abélard en Héloïse enkel geschreven zijn door Abélard (2020a).

7.1.1 Rowling

The Cuckoo's Calling is een misdaadverhaal dat verscheen in 2013 van de hand van Robert Galbraith. The Sunday Times werd echter getipt dat Galbraith een pseudoniem van Rowling was.

Op vraag van de krant liet Patrick Juola zijn computerprogramma los op *The Cuckoo's Calling*. Hij verzamelde teksten van Rowling en drie andere vrouwelijke misdaadschrijvers. Met *machine learning*-technieken kon hij achterhalen wie van de vier schrijfsters het meest waarschijnlijk de auteur van het boek was, Rowling dus.

Hoewel het onderzoek geen sluitend bewijs leverde, confronteerde The Sunday Times Rowling met het resultaat, waarop ze toegaf dat zij inderdaad het boek schreef onder een pseudoniem in de hoop onbevooroordeelde kritieken te ontvangen. Zonder de tip zou dit niet achterhaald zijn, aangezien het zonder enige aanwijzing zoeken is naar een speld in een hooiberg (Juola, 2013).

Een analyse die steunt op 4-grammen van karakters zou de tekst 'I LOVE YOU' opsplitsen in 'I LO', 'LOV', 'LOVE', 'OVE', 'VE Y', 'E YO', 'YOU' (Ainsworth & Juola, 2019).

The Cuckoo's Calling

Juola voerde testen uit waarbij werd gekeken naar woordlengte, woordenparen, 4-grammen van karakters en de honderd meest frequente woorden. De analyse toonde dat er stilistisch gelijkenissen zijn tussen *The Cuckoo's Calling* en ander werk van Rowling (*The Casual Vacancy*). Genoeg gelijkenissen om aan te nemen dat Rowling de auteur is (Juola, 2015; Ainsworth & Juola, 2019).

7.1.2 Caesar

De *Corpus Caesarianum* over de veldslagen van Julius Caesar bestaat uit vijf delen: *Bellum Gallicum*, *Bellum Civile*, *Bellum Alexandrinum*, *Bellum Africum* en *Bellum Hispaniense*.

Aulus Hirtius, een van Caesars generaals, beweerde in een brief aan Cornelius Balbus, ook een van Caesars aanhangers, dat hij meeschreef aan de *Corpus Caesarianum*, o.a. aan de *Bellum Gallicum* en de *Bellum Alexandrinum*. Hirtius stelde in de brief dat hij een stuk toevoegde aan de geschriften van Caesar over zijn veldtochten in Gallië, hiaten opvolde in de *Corpus Caesarianum* en er een slot aan breide.

Uit geschriften van o.a. Cicero weet men dat Caesar de eerste zeven boeken van de *Bellum Gallicum* zelf schreef. Men weet ook dat het einde van de *Bellum Hispaniense* ontbreekt. Er is ook geen grondige reden om te twijfelen aan de authenticiteit van de brief van Hirtius, maar de vraag blijft welke teksten van de *Corpus Caesarianum* Hirtius bijna 2000 jaar geleden juist schreef. Onderzoekers probeerden al meermaals om dit te achterhalen, maar de statistische methodes die ze voorhanden hadden noch het gebruik van computers boden uitsluitel.

De geboekte vooruitgang in de kunstmatige intelligentie, meer bepaald het succes van de neurale netwerken, forceerde een doorbraak: recent onderzoek toont dat er een duidelijk stilistisch verschil is tussen de eerste boeken van de *Bellum Gallicum*, geschreven door Caesar, en het laatste hoofdstuk ervan; en een groot deel van de *Bellum Alexandrinum* blijkt stilistisch sterk verwant met dat laatste hoofdstuk.

De onderzoekers schreven het laatste hoofdstuk van de *Bellum Gallicum* toe aan Hirtius en concludeerden dat Caesar behalve de eerste zeven boeken van de *Bellum Gallicum*, eveneens de *Bellum Civile* en de eerste 21 hoofdstukken van de *Bellum Alexandrinum* schreef;

Hirtius schreef het achtste, tevens laatste boek van de *Bellum Gallicum*, en de rest van de *Bellum Alexandrinum*. Noch Caesar noch Hirtius schreef de *Bellum Africum* en de *Bellum Hispaniense*, en deze zijn vermoedelijk ook niet van de hand van dezelfde auteur.

Deze conclusies liggen in de lijn van wat Hirtius in zijn brief schreef. Vermits het laatste deel van de *Bellum Hispaniense* ontbreekt, kan men bevestigen noch weerleggen dat Hirtius het slot schreef (Kestemont et al., 2016).

Corpus Caesarianum

De *Corpus Caesarianum* werd verdeeld in niet-overlappende stukken van 1000 woorden, die vervolgens als individuele documenten werden behandeld. Om stilistische gelijkenissen tussen de teksten op te sporen hebben Kestemont et al. ingezet op de woorden die voorkomen in de teksten, en op de n -grammen, omdat n -grammen in het bijzonder voor talen met verbuigingen, zoals het Latijn, hun nut hebben bewezen (2016).

7.1.3 *Abélard*

Jeroen De Gussem onderzocht met taalalgoritmes de schrijfstijl in de liefdesbrieven tussen Héloïse en Abélard (De Gussem, 2019, 2020a).

Deze briefwisseling in het Latijn is gesitueerd in Parijs in de 12de eeuw. Acht brieven vertellen het verhaal van de Parijse leraar Pierre Abélard (1079–1142) die een relatie begint met zijn leerling Héloïse. De geliefden worden echter uit elkaar gedreven en belanden uiteindelijk, gescheiden van elkaar, in het klooster.

Abélard en Héloïse hebben echt geleefd en waren zeker geliefden. Maar dat betekent niet dat deze brieven echt zijn, dat het hun privé-correspondentie is. Tot voor kort bleef dat een groot mysterie en het feit dat er geen originele brieven zijn, maakt het er niet gemakkelijker op.

Zie kader 'Abélard en Héloïse'; de functionele woorden zijn vetgedrukt. Figuur 7.2 geeft de gebruikte functionele woorden.

Abélard en Héloïse

Fragment uit de brieven: de originele Latijnse tekst (Troyes, Médiathèque) en een vertaling door Chris Tazelaar (Tazelaar, 1998).

ABELARD, eerste brief

- ... **quotiens mihi a** scholis reuerso uacaret, **tam in** die quam **in nocte ei** docendae operam darem, **et eam, si** negligentem sentirem, uehementer constringerem.

In qua re **quidem quanta eius** simplicitas esset uehementer admiratus, **non** minus **apud me** obstupui **quam si** agnam teneram famelico lupo committeret.

Qui cum eam mihi non solum docendam, **uerum etiam** uehementer constringendam traderet, **quid aliud** agebat **quam ut** uotis **meis** licentiam penitus daret, **et** occasionem, **etiam si** nollemus offerret **ut quam uidelicet** blanditiis **non** possem, minis **et** uerberibus facilius flecterem?

Sed duo erant **quae eum** maxime **a** turpi suspicione reuocabant, amor **uidelicet** neptis **et** continentiae **meae** fama praeterita.

Quid plura?

Primum domo **una** coniungimur, **postmodum** animo. **Sub** occasione **itaque** disciplinae amori penitus uacabamus **et** secretos recessus **quos** amor optabat studium lectionis offerebat.

Apertis **itaque** libris, plura **de** amore **quam de** lectione uerba **se** ingerebant; **plura** erant oscula **quam** sententiae; saepius **ad** sinus **quam ad** libros reducebantur manus; crebrius oculos amor **in se** reflectebat **quam** lectio **in** scripturam dirigebat. **Quoque** minus suspicionis haberemus, uerbera **quandoque** dabat amor, **non** furor, gratia, **non** ira, **quae** omnium unguentorum suauitatem transcenderent.

Quid denique?

Nullus a cupidis intermissus est gradus amoris **et, si quid** insolitum amor excogitare potuit, est additum.

Et quo minus **ista** fueramus experti gaudia, ardentius **illis** insistebamus, **et** minus **in** fastidium uertebantur.

Et quo me amplius **haec** uoluptas occupauerat, minus philosophiae uacare poteram **et** scholis operam dare.

- Hij [oom Fulbertus] verzocht mij, om telkens als ik na thuiskomst van mijn school vrije tijd had, mij met de taak onderwijs aan haar te geven te belasten, zowel overdag als 's nachts, en haar streng aan te pakken als ik merkte dat zij haar studie niet serieus opnam. Ik was buitengewoon verwonderd dat hij zo naïef was: als hij een weerloos lammetje had toevertrouwd aan een hongerige wolf, zou mijn verbazing niet groter geweest zijn. Immers, toen hij haar aan mij toevertrouwde, niet alleen om haar les te geven, maar ook om haar hardhandig de les te lezen, deed hij toch niets anders dan de volledige vrijheid geven aan mijn verlangens? Hij gaf me toch de gelegenheid, of ik die nu wenste of niet, om haar met standjes en klappen tot andere gedachten te brengen, als ik dat met lieve woordjes niet voor elkaar kreeg? Maar er waren twee krachtige argumenten die hem afhielden van boze vermoedens: zijn liefde voor zijn nichtje en de faam die ik verworven had vanwege mijn ingetogen leefwijze.

Wat zal ik er verder over zeggen? Eerst leefden wij in hetzelfde huis, daarna deelden wij er ons leven. De lessen waren de dekmantel waaronder wij ons volledig aan onze liefde overgaven, en de bestudering van teksten gaf ons de kans om ons af te zonderen voor de liefde. De boeken lagen open op tafel, maar we spraken meer over liefde dan over lectuur en we wisselden meer kussen dan meningen uit. Mijn handen belandden vaker bij haar boezem dan bij de bladzijden, en we lazen meer de liefde in elkaars ogen dan de regels in de boeken. Om maar niet de geringste verdenking te wekken gaf ik haar wel eens klappen, uit liefde, niet uit ergernis, en niet met boosheid maar met tederheid die de milde werking van alle balsems overtrof.

Wat zal ik verder nog zeggen? In onze wederzijdse begeerte bestegen wij alle trappen van genot, en als liefde daar nog iets ongewoons bij kon bedenken, dan gaven we ons ook daaraan over. En hoe minder ervaring wij met die vreugden hadden, des te vuriger stortten wij ons daarin en nooit raakten wij daarvan verzadigd.

HELOÏSE, tweede brief

- **Et si** uxoris nomen sanctius **ac** ualidius uidetur, dulcius **mihi semper** exstitit amice uocabulum; **aut, si non** indigneris, concubine **uel** scorti. **Ut quo me** uidelicet **pro te** amplius humiliarem, amplio**rem apud te** consequerer gratiam, **et sic etiam** excellentie **tue** gloriam minus lederem.

Deum testem inuoco, **si me** Augustus uniuerso presidens mundo matrimonii honore dignaretur, **totumque mihi** orbem confirmaret **in** perpetuo presidendum, **charius mihi et dignius** uideretur **tua** dici meretrix **quam illius** imperatrix.

- Al lijkt de aanduiding 'echtgenote' eerbaarder en duurzamer, voor mij klonk het woord 'vriendin' altijd zoeter, of zelfs, als je je daardoor niet gekwetst voelt, 'maîtresse' (concubine) of 'hoer' (scorti). Hoe meer ik me voor jou vernederde, des te meer hoopte ik jouw genegenheid te winnen zonder daarmee aan de glorie van jouw buitengewone persoonlijkheid afbreuk te doen [. . .] Ik roep God als mijn getuige aan dat ik, indien Augustus, die over de hele aarde heerste, mij de eer van een huwelijk waardig zou keuren en zou vaststellen dat de hele wereld voor eeuwig mijn bezit zou zijn, het voor mezelf verkieslijker en waardiger zou vinden jouw bijzit (meretrix) te heten dan zijn keizerin.



13de-eeuws handschrift van de brieven van Héloïse en Abélard (Troyes, Médiathèque, MS 802, 13e eeuw, ongekende oorsprong). Het deel uit de eerste brief van Abélard dat hierboven wordt geciteerd, is in geel aangeduid.

In de negentiende en twintigste eeuw werden reeds meerdere pogingen ondernomen om het auteurschap van de briefwisseling tussen Abélard en Héloïse te ach-

terhalen. De technieken waarover men toen beschikte, bleken daarvoor onvoldoende (De Gussem, 2019).

De Gussem heeft m.b.v. taalalgoritmes geanalyseerd hoeveel keer de meest frequente woorden en bepaalde functionele woorden, zoals bijwoorden, voegwoorden en lidwoorden, voorkomen in de brieven.

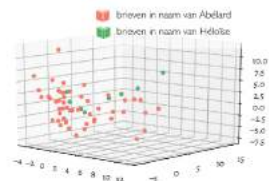
Deze frequenties kunnen grafisch worden voorgesteld, zoals te zien is in een vereenvoudigd voorbeeld in Figuur 7.1: de frequenties van drie woorden worden weergegeven door een punt dat geprojecteerd op de x-as, y-as en z-as respectievelijk de frequentie van het eerste, tweede en derde woord geeft. Op zo'n grafiek zullen de punten van teksten met dezelfde stijlkenmerken dicht bij elkaar liggen.

Op de grafieken in Figuur 7.3 wordt de clustering geïllustreerd voor de frequenties van drie woorden voor de briefwisseling tussen Héloïse en Abélard en ander werk van Abélard. Er wordt in getoond dat de briefwisseling qua stijl een sterke eenheid vormt, een stijl die gelijk is op die van Abélard.

In de praktijk gebeurt zo'n onderzoek niet met drie woorden, maar met heel veel woorden. Drie woorden zouden geen uitsluitsel over auteurschap kunnen geven. Een grafische voorstelling hiervan zou veel meer dan drie dimensies hebben en niet meer visueel te interpreteren zijn. Daarom wordt de meerdimensionale grafiek met wiskundige technieken, nl. *principal component analysis*, omgezet naar een twee- of driedimensionale, en dat op zo'n manier dat zoveel mogelijk informatie behouden blijft. Ook in het onderzoek van De Gussem gebeurde dat, zoals bv. te zien in Figuur 7.4. Ook op die grafiek is de scheiding tussen de auteurs mooi te zien.

Door het onderzoek van De Gussem ziet het ernaar uit dat de brieven geschreven zijn door een en dezelfde auteur, Abélard, die aan de hand van deze fictieve brieven

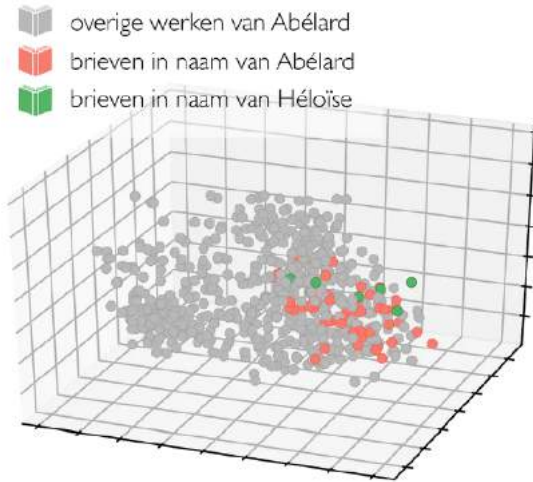
Bekijk de gebruikte functionele woorden in Figuur 7.2.



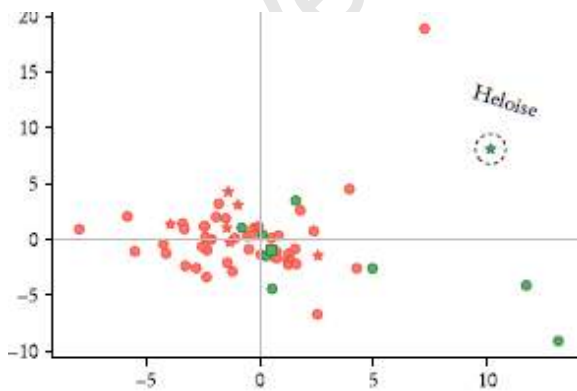
Figuur 7.1: Clustering van de frequentie van drie woorden voor de briefwisseling tussen Abélard en Héloïse. Slechts één cluster verschijnt (De Gussem, 2020a).

1-50	51-100	101-150	151-200	201-250	251-300	301-350
et	eum	ipsis	uestra	has	quamdam	alius
in	his	contra	tuam	quisquis	aliquem	ultra
non	ei	numquam	tui	utique	huc	utpote
est	illud	secundum	aliquo	qualis	meorum	illarum
ad	nisi	supra	meis	hos	deinde	neuter
ut	qua	hac	hunc	quocumque	intra	quoad
que	ipsum	modo	illos	rursum	ulla	proinde
de	ac	eas	quidam	illuc	nondum	etsi
quam	eam	illam	postea	quosdam	aliquibus	quodlibet
cum	ubi	illis	meo	tandem	quendam	qualibet
quod	scilicet	ante	ideo	semel	rursus	prope
qui	tamen	tua	diu	quamuis	uersus	uestros
si	tunc	aliquid	prout	tuas	quot	quoslibet
uel	ipso	mea	namque	talibus	quantus	utramque
etiam	earum	ipsam	extra	tum	quales	qualiter
hoc	apud	illo	aliud	aque	queque	quantumcumque
sed	ille	idem	quibusdam	plerisque	utriusque	unusquisque
autem	huius	usque	sola	aliquod	cuiusdam	qualem
nec	itaque	unus	quicumque	meos	rem	aliqui
eius	hanc	ergo	meus	talem	uestrum	istud
per	eorum	item	quidquid	utrum	at	quascumque
ab	super	tu	iis	donec	nuper	tuos
quo	ego	denique	aduersum	ue	aliquos	quiscumque
tam	post	aliqua	simul	quacumque	plerumque	propterea
quoque	solum	quodam	erga	quemadmodum	harum	quasdam
enim	illa	illius	eadem	utrumque	quibuscumque	cuiuslibet
sicut	nunc	quoniam	ipsorum	quatinus	alicuius	demum
se	uos	tuum	huic	uti	idcirco	hucusque
ex	sine	nullum	uestri	tuus	ipsarum	quicquam
quidem	ibi	unum	modum	mox	relique	tuorum
uero	tantum	postmodum	meam	totus	absque	utroque
ipse	eo	alia	horum	utraque	quodcumque	alicubi
aut	inter	sub	aliquam	alter	he	suus
nos	quantum	inde	coram	ob	quare	talia
atque	nulla	tamquam	aduersus	quandoque	procul	quamquam
quid	ipsius	quomodo	aliquando	ubique	uelut	plerosque
pro	hic	nam	utinam	alterum	quilibet	illic
quasi	hinc	quadam	uestram	etenim	is	aliquas
quia	eos	una	illorum	pleraque	talis	necnon
eis	immo	meum	quinque	recte	quisque	tuarum
ipsa	iuxta	illum	tuo	quoddam	siquidem	quousque
sic	propter	solus	ipsas	quotiens	quamcumque	interdum
quanto	ipsi	neque	umquam	hi	quantacumque	aliquarum
ita	illi	quando	olim	tantus	cuiuscumque	unicuique
id	seu	noster	quecumque	uester	quibuslibet	utriusque
unde	ea	licet	nullus	tot	meas	cuilibet
iam	adhuc	primum	circa	meorum	ubicumque	qualibus
quis	siue	tuis	tales	altera	quolibet	aliquorum
quippe	dum	ipso	iste	tali	sin	aduersa
uidelicet	mei	ista	illas	uestris	quidquam	fas

Figuur 7.2: Functionele woorden uit de brieven van Abélard en Héloïse (De Gussem, 2019).



Figuur 7.3: Vereenvoudigde voorstelling die toont dat de stijl van de brieven grote gelijkenissen vertoont met overige werken van Abélard (De Gussem, 2020a).



Figuur 7.4: Het werk waarvan geweten is dat het zeker van Héloïse is, ligt niet in de cluster die gevormd wordt door de werken van Abélard en de briefwisseling (De Gussem, 2019).

hun noodlottig liefdesverhaal de geschiedenis inschreef.

7.2 Plagiaat

Plagiaat checkers kunnen met ML plagiaat detecteren van bronnen die niet opgenomen zijn in een database, zoals bronnen in een vreemde taal en oudere bronnen die niet gedigitaliseerd zijn.

Men kan met NLP ook een vermoeden van *ghost writing* staven, tenminste als men voldoende teksten heeft van de persoon die beweert de auteur te zijn van een schrijfsel, maar het schrijven in werkelijkheid heeft uitbesteed (Juola, 2017).

7.3 Profilering

Auteursherkenning kan ook een rol spelen in een context van criminaliteit (Chaski, 2005; Ainsworth & Juola, 2019):

- Zo schreef iemand e-mails naar zijn baas waarin de baas racistisch werd bejegend. Nadien beweerde hij dat iemand anders de e-mails vanaf zijn computer had gestuurd zonder zijn medeweten.

Uit stilistisch onderzoek bleek dat hij ze wel degelijk zelf had geschreven.

- Een jonge, gezonde man werd door zijn *roommate* dood aangetroffen in bed. Uit de autopsie bleek dat hij een injectie had gekregen, die zijn dood had veroorzaakt. De *roommate* vond een zelfmoordbriefje op de gemeenschappelijke computer.

Uit stilistisch onderzoek bleek dat de kamergenoot het briefje had geschreven. Hij werd gearresteerd op vermoeden van het toedienen van de injectie, weliswaar om zijn vriend die ziek was te helpen, maar met een te hoge dosis medicatie.

Zo'n stilistisch onderzoek dat met de computer gebeurt en dat o.a. naar zinslengte, frequentie van woorden kijkt, heet forensische stylometrie.

- De *Unabomber* werd gevat door taalanalyse toe te passen op zijn manifest en brieven die hij schreef (FBI). Zie kader 'Unabomber'.

Unabomber

In 1978 ontplofte een *homemade* bom op de universiteit van Chicago. Meerdere bommen volgden, waarvan enkele in de vorm van een bombrief.

Een groot team onderzocht elke ontploffing minutieus. Maar omdat de bommenlegger de bommen met alledaags materiaal in elkaar stak en omdat hij zijn doelwitten willekeurig koos, had het team geen enkel bruikbaar spoor om hem te kunnen identificeren.

Er kwam een doorbraak in 1995. De bommenlegger bezorgde de FBI een manifest waarin hij zijn motieven uit de doeken deed. Het manifest werd gepubliceerd in *The Washington Post*, waarop het publiek mogelijke verdachten aandroeg.

David Kaczynski vermoedde dat zijn broer Theodore wel eens de dader zou kunnen zijn. Theodore had lesgegeven aan de universiteit van Berkeley waar twee bommen gelegd waren. David bezorgde de FBI brieven geschreven door zijn broer. Taalanalyse toonde aan dat deze papieren bijna zeker van dezelfde hand waren als het manifest.

Dit legde de basis voor zijn arrestatie.

De FBI deed er bijna twee decennia over om Theodore Kaczynski te pakken te krijgen.

(FBI).

7.4 Informatiebubbel

Algoritmes personaliseren de playlist die Spotify aanraadt, de films die Netflix aanbeveelt, en de reclame die wordt getoond op Google. Maar hoe zit het met nieuwsberichten?

Het onderzoek van Joris verschaft ons inzicht in de vaart die de zogenaamde *filter bubble* neemt (2021).

Informatiebubblel

Sinds de opmars van de sociale media komt het nieuws niet meer alleen langs de geschreven pers, televisie en radio tot bij de mensen. Ook Facebook, Twitter en Google bijvoorbeeld spelen daar nu een grote rol in. Mensen krijgen nu ook via de sociale media informatie over maatschappelijke thema's zoals het klimaat, het onderwijs en de COVID19-pandemie. Dat betekent dat sociale media een grote invloed hebben op wat de burger bezighoudt, en dus ook op wat op de politieke agenda komt.

Het nieuws op sociale media wordt m.b.v. algoritmes sterk gepersonaliseerd. Uit iemands surf- en klikgedrag wordt afgeleid welk soort artikels iemand interessant vindt, waarna artikels van dezelfde strekking en met dezelfde thema's prominent op een gepersonaliseerde startpagina van het platform verschijnen. Er wordt daarbij ook gekeken naar het klikgedrag van mensen met gelijksoortige interesses.

De algoritmes kennen daarvoor aan elk nieuwsbericht een score toe die weergeeft hoe dicht het bericht aansluit bij iemands interesses, en dus hoe groot de kans is dat iemand een bericht zal willen lezen. Om het thema en de inhoud van een artikel te bepalen, wordt een beroep gedaan op *Natural Language Processing*.

Socialemediaplatformen en nieuwsorganisaties kijken op een andere manier naar het aanbieden van nieuws. Socialemediaplatformen halen hun inkomsten vooral uit reclame. Nieuwsorganisaties hebben een maatschappelijke rol te vervullen, denk maar aan de VRT die een openbare omroep is.

Elk mediabedrijf doet zijn uiterste best om de lezer aan zich te binden. Behalve het personaliseren via nieuwsbrieven waarbij een gebruiker expliciet aangeeft wat zijn voorkeuren zijn, personaliseren ook nieuwsorganisaties al impliciet. De online platformen van de 'klassieke' media doen immers steeds meer een beroep op de aanbevelingsalgoritmes en *click bate*-berichten.

Voor de democratie is het belangrijk dat de burgers geïnformeerd zijn over de verschillende standpunten die er heersen over een bepaald onderwerp, en dat ze inhoudelijk geïnformeerd zijn over maatschappelijk relevante thema's.

Volgens recent wetenschappelijk onderzoek is er nog geen onmiddellijk gevaar voor de zogenaamde informatiebubble, omdat de meeste mensen nog steeds de actualiteit volgen langs verschillende kanalen. Maar hoe meer het personaliseren zich doorzet, hoe reëler de *filter bubble* wordt. De maatschappij kan er alvast op anticiperen, wil ze de democratie veiligstellen.

Men zou kunnen inzetten op de ontwikkeling van een diversiteitsalgoritme, zodat de burgers ook online van divers nieuws worden voorzien. Een diversiteitsalgoritme biedt artikels aan waarvan de inhoud verschilt met wat de lezer reeds aanklikte. Zowel in Nederland, Engeland als België wordt reeds onderzocht hoe een dergelijk algoritme zou kunnen functioneren.

Het effect van zo'n algoritme mag echter niet overschat worden. Er is een overvloed aan nieuwsartikels, nieuws komt steeds sneller online, en er is gewoonweg geen tijd om alles te lezen. Men mag daarbij ook niet vergeten dat mensen eerder artikels aanklikken die in de lijn liggen met hun interesses en waarin ze hun eigen opvattingen bevestigd zien. Mensen vinden de algoritmes die artikels aanbevelen aansluitend bij hun eerder leesgedrag, best wel handig. De lezers zullen dus aangemoedigd moeten worden om de artikels aanbevolen door een diversiteitsalgoritme toch te gaan lezen.

Om een aanbevelingsalgoritme of een diversiteitsalgoritme te trainen, heeft men gelabelde data nodig. Het label moet weergeven tot welke categorie een artikel hoort, bijvoorbeeld sport, politiek of klimaat, politiek links of rechts. Dit labelen gebeurt deels door mensen, maar door de grote toevloed aan nieuwe berichten, wordt dit proces zoveel mogelijk geautomatiseerd. Men traint dus ook NLP-algoritmes die nieuwe artikels kunnen labelen. Door de complexiteit van taal en taalbegrip is dit echter een zeer moeilijke opdracht.

(Joris, 2021; Lieman, 2019).

7.5 Persoonskenmerken

Taalgebruik hangt samen met geslacht, leeftijd, en persoonskenmerken zoals introvert of extravert zijn (Schwartz et al., 2013; De Gussem & Daelemans, 2020).

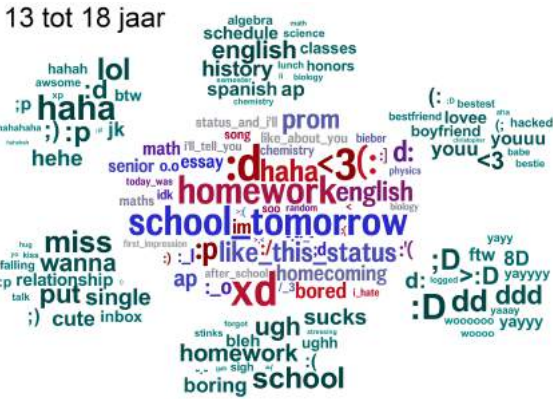
Zo zijn sommige woorden en woordengroepen meer typerend voor bepaalde leeftijdsgroepen, zoals 'huiswerk' voor tieners en 'familie en vrienden' voor volwassenen, en zijn er specifieke onderwerpen waarover men op Facebook schrijft, zoals 'familie' bij vrouwen en 'xbox' bij mannen. Extraverte personen gebruiken langere woorden dan introverte personen, en veel adjectieven.

Schwartz et al. illustreren dit met enkele woorden-

Persoonlijkheid kan worden bekeken volgens het Vijf Factor Model, waarbij de kenmerken 'extravert', 'aangenaam', 'zorgvuldig', 'neurotisch' en 'open voor ervaring' beschouwd worden (Schwartz et al., 2013).

wolken die ze hebben opgesteld a.d.h.v. Facebookposts die ze hebben geanalyseerd (zie Figuren 7.5, 7.6, 7.7, 7.8 en 7.9). De cluster in het midden bevat de woorden die typerend zijn voor de groep: hoe groter het woord, hoe meer typerend; hoe donkerder het woord, hoe frequenter. De clusters errond tonen de typerende onderwerpen van Facebookberichten.

Schwartz et al. gebruikten een dataset van 15 miljoen Facebookberichten van 75 000 vrijwilligers.



Figuur 7.5: Meest typerende woorden en onderwerpen voor tieners op Facebook (Schwartz et al., 2013).



Figuur 7.6: Meest typerende woorden en onderwerpen voor volwassenen op Facebook (Schwartz et al., 2013).

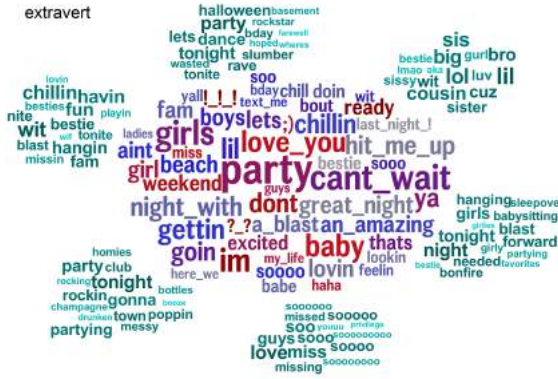
Het is dus niet verwonderlijk dat bedrijven socialemidiaposts bestuderen om een idee te krijgen van iemands persoonlijkheid om vervolgens de auteur te proberen te overtuigen om tot een bepaalde aankoop



Figuur 7.7: Meest typende woorden en onderwerpen voor vrouwen op Facebook (Schwartz et al., 2013).



Figuur 7.8: Meest typende woorden en onderwerpen voor mannen op Facebook (Schwartz et al., 2013).



Figuur 7.9: Meest typende woorden en onderwerpen voor extraverten op Facebook (Schwartz et al., 2013).

over te gaan, gebruikmakend van die informatie. Het is een manier om reclame te personaliseren.

7.6 *Human resource*

Zal er in de toekomst nog op de klassieke manier gesolliciteerd worden? Waarschijnlijk niet. Solliciteren is eigenlijk al grondig veranderd toen werkgevers toekomstige werknemers zijn beginnen te screenen op sociale media.

Maar het gaat al veel verder dan dat. Unilever bijvoorbeeld, start een sollicitatieprocedure met het verplicht indienen van een filmpje, wat dan door een AI-systeem wordt geanalyseerd. Men zegt dat het systeem er de meest gemotiveerde en geschikte kandidaten voor de job uithaalt (Oostra, 2019; Wall Street Journal, 2018).

Bekijk de video 'The Robots Are Now Hiring' (zie Kijktip).

Het Gentse bedrijf Traicie werkt aan een AI-systeem dat op basis van sollicitatiebrieven beslist in welke mate de auteur over de vaardigheden beschikt vereist voor een bepaalde job. Het systeem leidt daartoe persoonlijkheidskenmerken af uit wat de sollicitant schreef en zegt daarvoor een *unbiased* systeem te gebruiken.

Voor human resources is de persoonlijkheid van een sollicitant een interessant gegeven, sommige types zijn immers beter geschikt voor een bepaalde functie dan andere (De Gussem & Daelemans, 2020).

Discussie

Bekijk de video (De Winne, 2020):

<https://kanaalz.knack.be/nieuws/taalgebruik-op-cv-onthult-persoonlijkheid/video-normal-1678309.html>.

Bezoek ook de website <https://traicie.com/> en bespreek de objectiviteit van het systeem.

Discussie

In de volgende video worden de nieuwe tendensen bij het aanwerven van personeel aangehaald (Wall Street Journal, 2018):

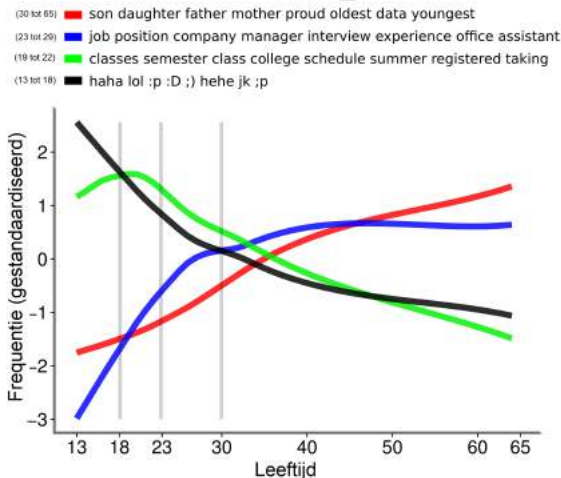
<https://youtu.be/8QEK7B9GUhM>

Bespreek.

7.7 Leeftijd inschatten

Instagram probeert m.b.v. AI de werkelijke leeftijd van gebruikers te schatten. Instagram wil immers het sturen van berichten door meerderjarigen aan minderjarigen onmogelijk maken, als die meerderjarige niet tot het Instagramnetwerk van de minderjarige behoort. Uit iemands taalgebruik kan men inderdaad zaken afleiden die naar een bepaalde leeftijdscategorie wijzen, zie Figuur 7.10 en de figuren uit paragraaf 7.5 (Schwartz et al., 2013).

Zie ook paragraaf 7.5



Figuur 7.10: Gebruik woorden naargelang de leeftijd (Schwartz et al., 2013). De grafiek toont dat de frequentie van sommige groepen van woorden daalt met de leeftijd, terwijl het gebruik van andere groepen toeneemt.

Samengevat

Computerwetenschappers ontwikkelen algoritmes die de schrijfstijl van een tekst onderzoeken. Elke auteur heeft een unieke schrijfstijl die hij in zijn teksten blootgeeft. Door de algoritmes in te zetten op een tekst kan de auteur misschien via de schrijfstijl achterhaald worden.

Deze taalalgoritmes voor auteursherkenning worden in verschillende domeinen toegepast, gaande van kunstgeschiedenis, over plagiaatdetectie tot antiterro-risme (Kestemont et al., 2016).

Voor de democratie is het belangrijk dat de burgers geïnformeerd zijn over de verschillende standpunten die er heersen over een bepaald onderwerp, en dat ze inhoudelijk geïnformeerd zijn over maatschappelijk relevante thema's.

Elk mediabedrijf doet zijn uiterste best om de lezer aan zich te binden, o.a. door nieuws te personaliseren met aanbevelingsalgoritmes. Volgens recent wetenschappelijk onderzoek is er nog geen onmiddellijk gevaar voor de zogenaamde informatiebubbel, omdat de meeste mensen nog steeds de actualiteit opvolgen langs verschillende kanalen. Maar hoe meer het personaliseren zich doorzet, hoe reëler de *filter bubble* wordt. Men kan er alvast op anticiperen met de ontwikkeling van een diversiteitsalgoritme, zodat de burgers ook online van divers nieuws worden voorzien.

Taalgebruik hangt samen met geslacht, leeftijd, en persoonskenmerken zoals introvert of extravert zijn. Het is dus niet verwonderlijk dat bedrijven socialemediaposts bestuderen om een idee te krijgen van iemands persoonlijkheid om vervolgens over te gaan tot het personaliseren van reclame of op zoek te gaan naar de meest geschikte kandidaat voor een job.

Kijktip

Hoe taalalgoritmes een middeleeuwse liefdesaffaire bekoelden (De Gussem, 2020b).

<https://www.youtube.com/watch?v=gI7a0yn1mw8>

Artificial Intelligence: The Robots Are Now Hiring (Wall Street Journal, 2018).

<https://youtu.be/8QEK7B9GUHM>

GESPROKEN TAAL

8.1 *Fysieke robots en digitale assistenten*

Om via gesproken taal over en weer te kunnen communiceren met een fysieke robot is zowel spraakherkenningstechnologie (spraak naar tekst) als spraaksynthesoftware (tekst naar spraak) nodig. De spraaksynthesoftware is er om de robot te laten spreken, de spraakherkenningssoftware opdat hij de mens die tegen hem spreekt, zou begrijpen. Dat geldt ook voor digitale assistenten zoals Siri op een iPhone, Alexa thuis en de *voice assistant* in de auto.

Bij deze toepassingen moet men rekening houden met dezelfde principes als bij een chatbot, behalve dat de conversatie via gesproken i.p.v. geschreven tekst verloopt en aanbieders van deze assistenten nog veel meer over de gebruiker kunnen te weten komen. De gesproken taal van de gebruiker wordt automatisch omgezet naar geschreven tekst via spraakherkenningstechnologie (spraak naar tekst) zodat de digitale assistent ermee aan de slag kan.

Alle technieken op getypte tekst, zoals sentimentanalyse, kunnen, zij het onrechtstreeks, ook toegepast worden op gesproken tekst. Het principe is net hetzelfde want spraaktekst wordt eerst omgezet naar geschreven

Er wordt onderzoek gedaan om bij deze toepassingen het omzetten van spraak naar tekst en omgekeerd te kunnen overslaan. Men beschikt al over experimentele, neurale netwerken die vertrekend van het stemgeluid direct een gesproken antwoord kunnen genereren. Zie de kader 'Real-time 'spraak naar spraak'-vertaling'.

tekst en daarop wordt sentimentanalyse uitgevoerd, net zoals je dat rechtstreeks op geschreven tekst zou doen.

Discussie

Start een discussie op over privacy bij gebruik van digitale assistenten aan de hand van deze tekst (Kenniscentrum Data & Maatschappij, 2021b):
<https://data-en-maatschappij.ai/publicaties/brainfood-databescherming-en-voice-assistants-1>

8.2 *Spraak naar tekst, tekst naar spraak*

Een aan een smartphone gedicteerde tekst wordt omgezet in geschreven tekst. De GPS in de auto spreekt de bestuurder toe om de weg te wijzen.

De AI-systemen die hiervoor gebruikt worden, kunnen geluid verwerken. Spraakherkenningstechnologie werkt met een opname van geluid die wordt omgezet naar een digitaal bestand. Het AI-systeem analyseert dan dat bestand. Dat is zeker geen eenvoudige taak aangezien iedereen zich op zijn eigen manier uitdrukt, zelfs in dezelfde taal. Bovendien moet het systeem menselijke spraak kunnen onderscheiden van eventueel achtergrondlawaai. ML-systemen beginnen echter ook al dialecten van eenzelfde taal van elkaar te kunnen onderscheiden.

Om een digitale assistent of robot te laten spreken, wordt een tekstbestand met spraaksynthesesoftware omgezet naar geluid (zie ook de kader 'Realistisch stemgeluid').

Ook voor de robot Alex in de toneelvoorstelling van Improbotics worden beide technologieën gebruikt. De gesproken tekst van de acteur wordt omgezet in getypte tekst, zodat het AI-systeem GPT-2 ermee aan de slag kan. Nadat de tekst van Alex door het AI-systeem werd gegenereerd, wordt deze opnieuw omgezet naar gesproken tekst. Zie ook de kader 'Improvistietheater met GPT-2' in hoofdstuk 3.

Realistisch stemgeluid

Als een AI-systeem spreekt, is het niet aangenaam als de stem weinig intonatie vertoont. Er is een specifiek onderzoeksdomein ontstaan dat zich hiermee bezighoudt. De Belg Aäron van den Oord, oud-student van het UGent IDLab, werkt bij Google Deepmind en verdiept zich hierin. Samen met zijn collega's bouwde hij een AI-model dat stemmen genereert die realistisch klinken. Improbatics gebruikt deze technologie om de robot Alex goed te laten klinken (van den Oord, 2017).

Real-time 'spraak naar spraak'-vertaling

'Spraak naar spraak'-vertalingen hebben als doel dat mensen met een verschillende moedertaal gemakkelijk met elkaar kunnen spreken.

Zulke systemen bestaan doorgaans uit drie componenten: spraakherkenning voor het omzetten van spraak naar tekst, een machinale vertaling van de tekst in de brontaal naar de doeltaal, en spraaksynthese voor het omzetten van de vertaling naar gesproken taal. Deze systemen werken al vrij goed, maar er zit wel wat vertraging tussen de originele spraak en de doelspraak. Echt vlotte conversaties zijn daardoor niet mogelijk.

Men streeft daarom naar snellere systemen door de omzettingen van spraak naar tekst en van tekst naar spraak te elimineren, en zo direct spraak in de ene taal te vertalen naar spraak in de andere taal. Google ontwikkelde Translatotron 2, een AI-systeem dat spectrogrammen als input neemt en dat spectrogrammen genereert van de vertaling in de doeltaal. Bovendien worden de kenmerken van de stem van de oorspronkelijke spreker behouden in de andere taal. Het komt er dus op neer dat een *deep learning*-model luistert naar iemand die spreekt in de ene taal, vertaalt wat er gezegd wordt naar de andere taal, en de vertaling laat uitspreken als ware de oorspronkelijke spreker vloeiend in deze andere taal. Google heeft het onmogelijk gemaakt om de vertaling te laten uitspreken met een andere stem; dit om te vermijden dat de tool zou gebruikt worden bij het maken van *deepfakes*.

(Jia & Weiss, 2019; Schwartz, 2021).

8.3 *Slim speelgoed*

Voor kinderen zijn er talloze ‘speelmaatjes’ op de markt, waar ze zich sterk kunnen aan gaan hechten. Deze sociale speelgoedrobots worden echter niet altijd goed onthaald. Met het ‘Internet of Toys’ worden kinderen betrokken in een commerciële, datagestuurde wereld, waar morele grenzen overschreden kunnen worden. Kinderen kunnen gevoelig zijn voor wat een robot hen vraagt of zegt te doen (Williams et al., 2018). Maar ook de privacy van de kinderen is in gevaar.

Mattel bracht Hello Barbie op de markt. Spraakherkenning zorgt ervoor dat het kind conversaties kan voeren met de pop. De pop, die verbonden is met servers in de cloud, leert uit vorige gesprekken. Omwille van privacyredenen kwam hier veel kritiek op en werd de pop van de markt gehaald.

Gesprekken tussen kinderen en hun teddybeer van CloudPets werden gehackt.

In Duitsland werd de slimme pop My Friend Cayla verboden omdat ze onder andere als afluisterapparatuur gebruikt kan worden door derden.

Robothond Aibo mag omwille van privacyredenen niet verkocht worden in de EU. Aibo neemt bijvoorbeeld foto’s van de woning waarin hij vertoeft en die worden opgeslagen in de cloud, zo komen ze bij de fabrikant terecht, net zoals de conversaties die met Aibo worden gevoerd.

Sociale robots zijn robots die kunnen communiceren met mensen. Ze worden in onze samenleving in uiteenlopende gedaantes ingezet in verschillende sectoren, zoals de zorgsector en de horeca.

8.4 *Zelfrijdende auto’s*

Iedereen is er al aan gewend dat een gps spreekt tegen de autobestuurder. In sommige auto’s kan de bestuurder via een *voice assistant* de airco aanzetten, iemand opbellen of boodschappen bestellen. Dat verhoogt mogelijk de veiligheid want de bestuurder hoeft het stuur niet los te laten en kan zijn ogen op de weg houden.

Het is echter ook de bedoeling dat een bestuurder

tegen de zelfrijdende auto zal kunnen spreken om instructies te geven. Om vlot te kunnen communiceren met de bestuurder is er dus best een *voice assistant* aanwezig die zowel over spraakherkennings- als spraak-synthesoftware beschikt.

Er is nog veel werk aan de winkel. Stembesturing zal goed moeten werken op alle soorten stemmen; het systeem moet achtergrondgeluid van bv. de motor en muziek kunnen onderscheiden van een stem; het systeem moet ook goed kunnen omgaan met dialecten en accenten (Walch, 2020). Volgens een rapport van het Capgemini Research Institute uit 2019 overschatten autoproducenten de capaciteiten van de *in-car voice assistants* (zie Figuur 8.1). Bijna 60 % van de gebruikers vindt dat de *in-car voice assistant* aan verbetering toe is. Ze vinden bv. dat de assistent hen niet altijd goed begrijpt. Sommigen maken zich zorgen om hun privacy.

De zelfrijdende wagen zal een gegeven instructie, zoals “Parkeer naast de rode auto”, niet ‘blindelings’ mogen uitvoeren. Hij zal daarbij rekening moeten houden met de fysieke wereld waarin hij veel objecten moet herkennen. Daar is dus beeldherkenningssoftware voor nodig.

Er komen heel wat technologieën kijken bij zo’n autonome wagen, bijvoorbeeld het efficiënt inzetten van afstandssensoren, camera’s en ruisonderdrukkende microfoons, en het beschikken over een robuuste internetverbinding.

Er zijn ook ethische aspecten die aandacht verdienen. Hoe worden de data die de systemen verzamelen, bewaard, gebruikt, gedeeld en verwijderd?

Het eerder vermelde rapport van het Capgemini Research Institute geeft ook aan dat een goede samenwerking tussen de *in-car voice assistant*, de *home assistant* en de digitale assistent op de smartphone van de gebruikers gewenst is. Dan zouden ze bv. met de *in-car*



Figuur 8.1: Percentage van de leidinggevenden in de auto-industrie (donker-groen) en van de gebruikers (lichtgroen) dat gelooft dat de *voice assistant* in de auto tegemoetkomt aan de wensen en voorkeuren van de gebruiker. Het rapport is het resultaat van een bevraging bij meer dan 7000 gebruikers en 300 leidinggevend in de auto-industrie. Er werden ook meer dan 10 diepte-interviews afgenomen van leidinggevend en *voice-service providers* (Capgemini Research Institute, 2019).

voice assistant de garagepoort kunnen openen en de verwarming thuis kunnen aanzetten. Maar zo'n integratie brengt ook aspecten van privacy met zich mee: er worden dan immers data gedeeld tussen meerdere toepassingen (Walch, 2020).

Samengevat

Om via gesproken taal over en weer te kunnen communiceren met een fysieke robot is zowel spraakherkenningstechnologie (spraak naar tekst) als spraaksynthesesoftware (tekst naar spraak) nodig. De spraaksynthesesoftware is er om de robot te laten spreken, de spraakherkenningssoftware opdat hij de mens die tegen hem spreekt, zou begrijpen.

Alle technieken op getypte tekst, zoals sentimentanalyse, kunnen, zij het onrechtstreeks, ook toegepast worden op gesproken tekst. Het principe is net hetzelfde want spraaktekst wordt eerst omgezet naar geschreven tekst (spraak naar tekst) en daarop wordt sentimentanalyse uitgevoerd, net zoals je dat rechtstreeks op geschreven tekst zou doen.

Spraakherkenning en spraaksynthese worden bijvoorbeeld toegepast bij slim speelgoed, sociale robots, 'spraak naar spraak'-vertaling en stembesturing.

Kijktip

Alexa after man goes to dentist (Shatz, 2018).

<https://www.youtube.com/watch?v=IKZToY-V16w>

Demo: The magic of AI neural TTS and holograms at Microsoft Inspire 2019 (Microsoft, 2019).

<https://youtu.be/auJJrHgG9Mc>

We Interviewed Pepper - The Humanoid Robot (Tech Insider, 2018).

<https://www.youtube.com/watch?v=zJHyaD1psMc>

IBM Project Debater (IntelligenceSquared Debates, 2019).

https://youtu.be/3_yy0dnIc58

9.1 *Nieuwe technologieën*

Nieuwe technologieën veranderen de maatschappij en sommige doen dat op een zeer ingrijpende manier. Volgens Gabriels (2019) "geven technologische ontwikkelingen mee vorm aan hoe wij de wereld zien en beïnvloedt technologie ons denken, ons gedrag en onze waarden en normen".

De uitvinding van het elektrische licht heeft geleid tot de bouw van elektriciteitscentrales en heeft er uiteindelijk voor gezorgd dat de productie in fabrieken sterk werd opgedreven. Nu kan onze maatschappij niet meer functioneren zonder elektriciteit. Met de komst van de wasmachine moest minder tijd besteed worden aan het wassen van kleren, maar we zijn ze ook meer gaan wassen. Door echografie te gebruiken, gaan we anders om met een zwangerschap en eventuele complicaties die zich (kunnen) voordoen.

Het internet en de alomtegenwoordige smartphone hebben een grote invloed op de arbeidsmarkt, ons huishouden en de vrijetijdsbesteding, m.a.w. op ons dagdagelijks leven. Communicatietoepassingen en sociale media hebben de dienstverlening van de overheid en de bedrijven veranderd. Online shoppen, het massaal

delen van foto's, digitaal bankieren, een recept of een goedkopere energieleverancier opzoeken via het internet, een hotel boeken vanuit je luie zetel... Deze zaken zijn zo goed ingeburgerd dat een leven zónder haast ondenkbaar is.

Hoe ingrijpend een nieuwe technologie de maatschappij zal veranderen, is moeilijk in te schatten. De uitvinding van de telefonie en van de fotografie in de negentiende eeuw hebben bepaald hoe de hedendaagse communicatie eruitziet. Toen de eerste auto's begonnen te rijden, had niemand verwacht dat men ooit met zijn allen uren in de file zouden staan. Wie kon in 1969 vermoeden dat we nu met een smartphone op zak zouden lopen met meer rekenkracht dan de computers die de astronauten naar de maan brachten? In 1995 was het moeilijk in te schatten dat het internet vandaag zo'n grote impact zou hebben op het dagelijkse leven.

Het is echter niet alleen moeilijk om de impact van een nieuw technologisch product in te schatten als het nog niet in gebruik is, eens wijdverspreid is het ook moeilijk om aanwezige ongewenste effecten te corrigeren (Collingridge dilemma) (Gabriels, 2019).

Naast de directe impact van nieuwe technologieën is er ook een indirecte impact waarvoor men bedachtzaam moet zijn (Eynikel, 2017). Ook 'positieve' technologieën hebben soms een minder zichtbaar negatief effect.

Er zijn hiervan tig voorbeelden, zoals de industriële revolutie die een stijging van atmosferische broeikasgassen veroorzaakte, het gegeven dat een sonarboot walvissen in de war brengt en robotmaaiers die egels verwonden. Dankzij de moderne communicatie is men sneller op de hoogte van wat er in de wereld gebeurt, maar de keerzijde hiervan is dat men voortdurend op zijn hoede moet zijn voor *fake news*.

9.2 AI-technologie

Discussie. Moeder spreekt met overleden kind.

Nayeon, het dochtertje van de Zuid-Koreaanse Jang Ji-sung, stierf in 2016 aan een ongeneeslijke ziekte. Om de mogelijkheden van virtual reality (VR) te demonstreren werd in 2020 met VR een reünie tussen Nayeon en Jang gesimuleerd en uitgezonden op televisie. Om de virtuele reünie tot stand te brengen maakte men gebruik van VR, spraaktechnologie en spraaksynthese a.d.h.v. video's van Nayeon (Ng, 2020; Bergan, 2020; Grossman, 2020; Kim, 2020).

MBClife. I met you. <https://youtu.be/uf1TK8c4w0c>
<https://interestingengineering.com/vr-reunites-mother-with-deceased-daughter>
<https://www.independent.co.uk/news/world/asia/dead-child-virtual-reality-mother-daughter-south-korea-jang-ji-sung-a9325691.html>
<https://slate.com/technology/2020/05/meeting-you-virtual-reality-documentary-mbc.html>
<https://www.inverse.com/innovation/korean-vr-reunites-family-beyond-grave>

Als men de AI-technologie goed inzet, kan die bijdragen tot een betere wereld. Denk maar aan exoskeletten die mensen met een beperking meer mobiel maken, ouderen die langer thuis kunnen blijven wonen, een meer efficiënte voedselproductie, zuinigere elektrische toestellen, veiliger werk door robots in te zetten bij de ontmijningsdienst. Maar zal AI te allen tijde worden ingezet voor een mooiere maatschappij? Bij de ontwikkeling van AI-systemen kan men zich door verschillende motieven laten leiden. Zal men bij de ontwikkeling van een nieuw product steeds de mens centraal zetten, of soms ook het AI-systeem zelf of geldbejag?

Door robots in te zetten in de zorgsector krijgen verplegers meer tijd voor de patiënten. Als men er echter voor kiest om minder personeel aan te werven, dan verliest men de mens uit het oog. Het inzetten van technologie in de geneeskunde kan een positief iets zijn, maar hoe ver wil men daarin gaan? Welk soort cyborg is toelaatbaar? Is een virtuele therapeut zoals Woebot wenselijk? Welke taken krijgen zorgrobots? Hoe bekomt men dat exoskeletten voor iedereen toegankelijk zijn en niet enkel voor mensen met voldoende geld (Gabriels, 2019)?

Discussie

Welk soort cyborg is toelaatbaar?

Is een virtuele therapeut zoals Woebot wenselijk?

Welke taken krijgen zorgrobots?

Hoe bekomt men dat exoskeletten voor iedereen toegankelijk zijn en niet enkel voor mensen met voldoende geld?

Onder het mom van het bestrijden van de criminaliteit kan men AI-systemen inzetten om mensen overal gade te slaan. De privacy van mensen komt dan wel in het gedrang.

Een routeplanner doet ook suggesties voor restaurants in de buurt van het traject. Hoe wordt er bepaald welke restaurants er worden getoond en welke motieven spelen daarin mee?

Discussie

Burgers geven privacy op in ruil voor meer veiligheid. Gebruikers geven privacy op in ruil voor een dienst.

Bespreek.

9.3 Bias

Discussie

Home assistants en *smartphone assistants* hebben vaak een vrouwelijke stem. Sociale robots die worden ingezet aan de balie van een hotel krijgen een vrouwelijk uitzicht en een vrouwelijke stem. Deze praktijk blijkt niet onschuldig te zijn, want het weerspiegelt de vooroordelen van het team dat het AI-systeem ontwikkelde, m.a.w. vooroordelen die aanwezig zijn in de maatschappij: vrouwen worden aangezien als zorgzaam en worden in een dienstverlenende rol geduwd. Omdat digitale assistenten en sociale robots steeds meer gebruikt zullen worden, zouden deze vooroordelen kunnen worden versterkt. Het wordt dan extra moeilijk om stereotypen dat bv. zorgen en dienst verlenen typisch vrouwelijk zijn, te doorbreken (Chin & Robinson, 2020).

Bespreek.

Spraakherkenning werkt niet even goed op alle stemmen. Autofabrikanten kregen te horen dat de stembesturing beter mannenstemmen herkende dan vrouwenstemmen omdat de database van stemopnames om het systeem te trainen veel meer mannenstemmen bevatte. Ook door dialecten en etniciteit kan iemand anders klinken, waardoor het systeem niet goed werkt (Tatman, 2020).

Tijdens een schoonheidswedstrijd in 2016, gejureerd door een AI-systeem, werden deelnemers met een donkere huidskleur benadeeld, omdat het algoritme getraind was met foto's van vooral blanken.

In tegenstelling tot wat men doorgaans denkt, neemt een AI-systeem geen 100 % 'objectieve' beslissingen. Een AI-systeem is 'vooringenomen', meestal te wijten aan de data waarmee het systeem werd getraind. Deze bias moet men zoveel mogelijk controleren en tot het minimum herleiden (zie kader 'Bias').

Ontwikkelaars van AI-systemen moeten reeds bij de ontwikkeling bedacht zijn op mogelijk ongewenste effecten. De morele verantwoordelijkheid van een AI-

systeem dat in de maatschappij wordt gebracht, ligt niet bij het AI-systeem, maar bij de mens.

Bias

Bias komt voor als de data niet representatief zijn. Als men bv. enkel groene appels als voorbeelden geeft aan een DL-systeem, dan zal het model geen rode appels herkennen. Of als men bij de training enkel foto's van honden onder een stralend blauwe hemel aanbiedt, dan zal het DL-model een hond in de regen niet bij de klasse 'hond' indelen. Als de gebruikte data gekleurd zijn door een aanwezige bias in de maatschappij, zoals stereotypen, dan zal dit ook doorgegeven worden aan het ML-systeem. Men moet er dus over waken dat het model daardoor niet discriminerend wordt voor bepaalde bevolkingsgroepen. Als men bv. enkel vrouwelijke verplegers in de dataset stopt, dan zullen mannen niet als verpleger worden geclassificeerd. Een model wordt nochtans getest voor het in gebruik genomen wordt. De testdata kunnen echter dezelfde bias bevatten als de trainingdata.

Dankzij AI kan men op sommige platforms interessante vacatures in de mailbox krijgen. Maar krijgt men wel alle vacatures te zien waarin men geïnteresseerd is? Tot 2018 gebruikte Amazon een sterk bejubeld maar ondertussen afgevoerd AI-systeem om sollicitanten te beoordelen. Het systeem selecteerde geen vrouwen voor technologische posities. Het was immers getraind met historische data: sollicitanten van de voorbije 10 jaar, voornamelijk mannen aangezien zij nog steeds de technologiewereld domineren. Een Gents bedrijf werkt aan een AI-systeem dat uit sollicitatiebrieven beslist in welke mate de auteur geschikt is voor de job. Het systeem leidt daartoe persoonlijkheidskenmerken af uit wat de sollicitant schreef. Sommige bedrijven maken om nieuw personeel aan te werven, gebruik van een video-interview, dat geanalyseerd wordt door een algoritme. Daarbij wordt niet enkel rekening gehouden met wat de sollicitant zegt, maar ook met gelaatsuitdrukkingen bijvoorbeeld.

Discussie

Kijk naar Filmpje van Kanaal Z (De Winne, 2020):

<https://kanaalz.knack.be/nieuws/taalgebruik-op-cv-onthult-persoonlijkheid/video-normal-1678309.html>

Bespreek.

Bekijk de video van Furhat Robotics: Unbiased Recruiter Robot (Furhat Robotics, 2018).

<https://youtu.be/rPKrdxiEkQ0>

Bespreek.

Het nemen van beslissingen wordt steeds meer geautomatiseerd. Algoritmen bepalen welk nieuws je te zien krijgt op Facebook en welke filmpjes op YouTube worden geweigerd, maar ook de banken en de overheid maken er gebruik van. Daarom klinkt de vraag naar transparantie steeds luider. Men moet er zich echter ook bewust van zijn dat transparantie over het gebruikte algoritme nog niet betekent dat het ook betrouwbaar is. Men kan een AI-systeem niet moreel verantwoordelijk stellen voor zijn handelingen (Eynikel, 2017). Een mens neemt morele beslissingen door zorgvuldig af te wegen wat hij belangrijk vindt, rekening houdend met zijn normen en waarden. Hij zal daarbij ook beïnvloed worden door menselijke relaties, empathie. Zijn normen en waarden zijn deels bepaald door de cultuur waarin hij opgroeide of leeft. Dat betekent dat wat bij ons moreel aanvaard is, het daarom niet is in Japan en vice versa. Als men wil dat een slimme robot zo handelt dat een mens die handeling moreel verantwoord vindt, dan zal de mens daarvoor moeten zorgen. Dat betekent bv. dat hoe meer autonomie men toekent aan een AI-systeem, bv. een zelfrijdende auto, hoe groter de morele verantwoordelijkheid wordt van de mens, hetzij als ontwikkelaar of als producent, hetzij als wetgever, hetzij als gebruiker. Nu al moet er nagedacht worden

over de mogelijke impact die toekomstige ontwikkelingen in technologie op mens en maatschappij zullen hebben. Men moet nadenken over hoe men de systemen betrouwbaar kan maken en hoe men de wetten moet aanpassen. Men kan niet wachten met een wetgeving rond zelfrijdende auto's tot die auto's massaal de weg opgaan. Ethische kwesties vergen veel tijd. Ze zijn ook vaak complex. Er is in elk geval al aandacht voor: ethical AI , responsible AI , AI4good, trustworthy AI , explainable AI zijn termen die her en der opduiken, met de bijbehorende richtlijnen.

Zie ook paragraaf 8.4.

Seksistische vertaling

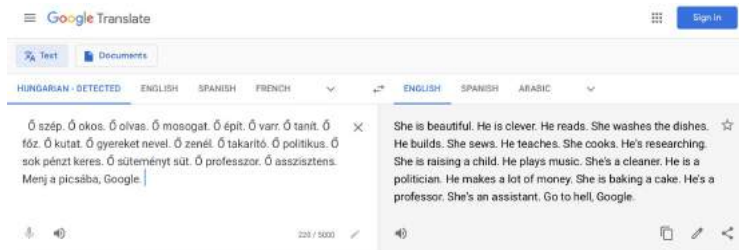
Bij gebruik van Google Translate kan men geconfronteerd worden met verkeerde, seksistische vertalingen. Illustreer dit door een Engelse tekst naar een andere taal, bv. het Nederlands, te vertalen en nadien terug naar het Engels.



Afbeelding uit een tweet van ZombieTron waarin een tekst wordt vertaald van het Engels naar het Hongaars en dan terug naar het Engels (2021).

Is Google Translate seksistisch?

In sommige talen, bv. Fins, Hongaars, Maleis, en Filipijns, is er slechts één voor-naamwoord voor de derde persoon enkelvoud. Ze maken dus geen onderscheid tussen hij, zij en het. Bij een vertaling naar pakweg het Engels, zal het vertaalprogramma dan een keuze maken, helaas vaak een seksistische. Dokters bv. worden aangezien als mannen en verplegers als vrouwen. Het AI-systeem achter Google Translate is getraind op heel veel voorbeelden waarin het AI-systeem op zoek gegaan is naar patronen. Wat het meest voorkwam in de voorbeelden, zal dan ook tot uiting komen in de vertalingen. Maatschappelijke vooroordelen en *gender bias* worden erin weerspiegeld.



Afbeelding uit een tweet van Dora Vargha waarin een vertaling wordt getoond van het Hongaars naar het Engels (2021).



Afbeelding uit een tweet van Maria Pettersson waarin een vertaling wordt getoond van het Fins naar het Engels (2021).

Discussie

Ga zelf op zoek naar voorbeelden van bias in AI-systemen.

Samengevat

Nieuwe technologieën veranderen de maatschappij en sommige doen dat op een zeer ingrijpende manier. De impact van AI zal mogelijk nog groter zijn dan de komst van het internet.

AI is al aanwezig in het dagelijks leven en is ook al doorgedrongen in de juridische wereld en de gezondheidszorg. AI speelt ook een rol bij het opsporen van kunstvervalsingen en geeft nieuwe onderzoeksdomeinen, zoals computerlinguïstiek, een boost.

Hoewel men optimistisch is over de toekomstige ontwikkelingen binnen AI, is enige realiteitszin op zijn plaats.

Men moet beseffen dat sommige zaken veel tijd zullen vergen, bv. vertalen tussen om het even welke talen en het begrijpen van taal.

Net zoals bij andere technologische ontwikkelingen steken ook hier ethische dilemma's de kop op. In tegenstelling tot wat men doorgaans denkt, neemt een AI-systeem geen 100 % 'objectieve' beslissingen. Een AI-systeem is 'voorin genomen', meestal te wijten aan de data waarmee het systeem werd getraind. Deze *bias* moet men zoveel mogelijk controleren en tot het minimum herleiden. Bias komt voor als de data niet representatief zijn. Als de gebruikte data gekleurd zijn door een aanwezige bias in de maatschappij, zoals stereotypen, dan zal dit ook doorgegeven worden aan het ML-systeem.

Kijktip

This 'Racist soap dispenser' at Facebook office does not work for black people. https://youtu.be/YJjv_0eiHmo (Futureism, 2017).

Leestip

Mark Coeckelbergh. AI ETHICS (Coeckelbergh, 2020).

Caroline Criado Perez. Onzichtbare vrouwen (Criado Perez, 2019).

Katleen Gabriels. Onlife (Gabriels, 2016).

Katleen Gabriels. Regels voor robots. Ethiek in tijden van AI (Gabriels, 2019).

Versie 1.0

DE NOTEBOOKS

10.1 *Python en Jupyter Notebook*

Notebooks zijn digitale documenten die zowel uitvoerbare code bevatten als tekst, afbeeldingen, video, hyperlinks ... Het is dus mogelijk om nieuwe begrippen aan te brengen, opdrachten te formuleren, foto's te tonen, bepaalde fundamenten van programmeren uit te leggen, aanwezige code uit te voeren en zelf code op te stellen.

Jupyter Notebook is een opensourcewebapplicatie waarin men notebooks kan creëren.

De notebooks bestaan eigenlijk uit een opeenvolging van cellen: Markdown-cellen voor tekst en beeld en code-cellen om te programmeren.

De notebooks van het project 'Chatbot' omvatten een notebook over de principes van programmeren in een context van de dt-regels, een notebook over het ontwikkelen van een chatbot en notebooks over sentimentanalyse.

De notebooks worden aangeboden via een webserver, toegankelijk via een webbrowser. Op de gebruikte computers moet geen extra software worden geïnstalleerd om met de notebooks aan de slag te gaan.



Figuur 10.1: De notebooks zijn online beschikbaar via <https://www.aiopschool.be/chatbot>.

In de notebooks worden gekleurde kaders gebruikt. De groene vertellen bv. wat de notebook behandelt en leggen linken naar de leerinhouden van het 'Chatbot'-project.

Het programmeren gebeurt met Python 3. Python is een zeer toegankelijke programmeertaal, die vaak ook zeer intuïtief is in gebruik. Bovendien is Python zo populair dat er heel wat modules voorhanden zijn die men vrij kan gebruiken. Python draagt zo ook bij tot de democratisering van deep learning. In een module zitten heel wat functies vervat die ervaren informatici reeds hebben geprogrammeerd. Wie dat wilt, kan die modules en de bijbehorende functionaliteiten gebruiken. Maar daarvoor moet men eerst de gewenste module importeren in het Python-script.

Python is een objectgeoriënteerde taal. Alles is er een object: elk getal, elk stuk tekst ... Een object neemt een bepaalde plaats in in het geheugen en heeft een bepaalde waarde. Elk object heeft ook een type, bv. integer, string en list. Alle objecten met eenzelfde type behoren tot dezelfde klasse. Binnen deze klasse zijn er methodes voorzien die men specifiek voor objecten van dit type kan oproepen. Naast deze methodes zijn er ook functies die men kan gebruiken, zoals `print()`, om iets te laten verschijnen op het scherm, en `input()`, om iets op te vragen aan de gebruiker.

10.2 Praktisch

Bij het begin van een notebook worden eerst de nodige modules geïmporteerd. In de 'Chatbot'-notebooks zijn dat modules om natuurlijke taal te kunnen verwerken.

Tekst wordt ingevoerd met Markdown in een Markdown-cel. Code wordt ingevoerd in een code-cel. Cellen kunnen worden bewerkt, verwijderd, toegevoegd, geko-

Een programma in Python noemt men een *script*.



De belangrijkste instructies en sneltoetsen om vlot met de notebooks te kunnen werken, vind je in de vier notebooks die het werken met de notebooks uitleggen.

pieerd, geknipt en geplakt.

Interessant om te weten is dat in een notebook alle code samen hoort. De notebook onthoudt als het ware welke code reeds werd uitgevoerd, ongeacht in welke volgorde die werd ingetikt in de notebook. Het is het tijdstip van uitvoeren dat de sequentie van het uiteindelijke script bepaalt.

We stimuleren dat er bij het invoeren van code aandacht wordt besteed aan een leesbare programmeerstijl en de nodige verduidelijkende commentaar.

In de eerste notebook van het 'Chatbot'-project worden fundamentele zaken aangebracht, zoals het invoeren van een printopdracht en het werken met strings en lijsten.

Samengevat

Notebooks zijn digitale documenten die zowel uitvoerbare code bevatten als tekst, afbeeldingen, video, hyperlinks ...

De notebooks bestaan uit een opeenvolging van cellen: Markdown-cellen voor tekst en beeld en code-cellen om te programmeren.

Het programmeren gebeurt met Python 3.

De notebooks van het project 'Chatbot' omvatten een notebook over de principes van programmeren in een context van de dt-regels, een notebook over het ontwikkelen van een chatbot en notebooks over sentimentanalyse.

EINDTERMEN

De leerkracht bepaalt zelf welke leerdoelen aan bod komen in het 'Chatbot'-project. Aan de eindtermen van 'Digitale competentie en mediawijsheid' voor de tweede en derde graad (zoals ze momenteel te vinden zijn op <https://onderwijsdoelen.be/>) kan zeker gewerkt worden.

De gelinkte eindtermen zijn hier opgenomen. De leerplandoelen zullen verschijnen op de webpagina <https://aiopschool.be/chatbot>.

11.1 Digitale competentie en mediawijsheid

11.1.1 Eindtermen digitale competentie en mediawijsheid

Eindtermen tweede graad

- *4.1 De leerlingen tonen zelfvertrouwen bij het verkennen en gebruiken van digitale infrastructuur en toepassingen (attitudinaal).*
- *4.2 De leerlingen gebruiken doelgericht en adequaat standaardfunctionaliteiten van digitale infrastructuur en toepassingen om digitaal inhoud te creëren, te delen en te beheren. (Toepassen).*
- *4.3 De leerlingen gebruiken doelgericht en adequaat standaardfunctionaliteiten van digitale infrastructuur en toepassingen om digitaal te communiceren, samen te werken en te participeren aan initiatieven. (Toepassen).*
- *4.4 De leerlingen lichten toe hoe bouwstenen van digitale systemen zich tot elkaar verhouden en op elkaar inwerken. (Begrijpen).*

- 4.5 (Finaliteit doorstroom) De leerlingen ontwerpen algoritmen om problemen digitaal op te lossen. (Creëren).
- 4.5 (Dubbele finaliteit) De leerlingen lossen een afgebakend probleem digitaal op door een aangereikt algoritme aan te passen. (Creëren).
- 4.6 (Finaliteit doorstroom en dubbele finaliteit) en 4.5 (Finaliteit arbeidsmarkt) De leerlingen leggen wederzijdse invloeden uit tussen enerzijds het individu en anderzijds media, digitale infrastructuur en digitale toepassingen. (Begrijpen).
- 4.7 (Finaliteit doorstroom en dubbele finaliteit) en 4.6 (Finaliteit arbeidsmarkt) De leerlingen passen regels van de digitale wereld toe. (Toepassen).
- 4.8 (Finaliteit doorstroom en dubbele finaliteit) en 4.7 (Finaliteit arbeidsmarkt) De leerlingen evalueren mogelijkheden en risico's van hun mediagedrag en dat van anderen. (Evalueren).

Eindtermen derde graad

- 4.1 De leerlingen tonen zelfvertrouwen bij het verkennen en gebruiken van digitale infrastructuur en toepassingen (attitudinaal).
- 4.2 De leerlingen gebruiken doelgericht en adequaat functionaliteiten van digitale infrastructuur en toepassingen om digitaal inhoud te creëren, te delen en te beheren. (Toepassen).
- 4.3 De leerlingen gebruiken doelgericht en adequaat functionaliteiten van digitale infrastructuur en toepassingen om digitaal te communiceren, samen te werken en te participeren aan initiatieven. (Toepassen).
- 4.4 De leerlingen beoordelen bouwstenen van digitale systemen in functie van het eigen gebruik en het gebruik ervan in een maatschappelijke context. (Evalueren).
- 4.5 (Finaliteit doorstroom) De leerlingen programmeren volgens een bepaalde systematiek oplossingen voor problemen aan de hand van zelfontworpen algoritmes. (Creëren).
- 4.5 (Dubbele finaliteit) De leerlingen lossen een complex probleem digitaal op door een aangereikt algoritme aan te passen. (Creëren).
- 4.6 (Finaliteit doorstroom en dubbele finaliteit) en 4.5 (Finaliteit arbeidsmarkt) De leerlingen stellen consequent passend mediagedrag (attitudinaal).
- 4.7 (Finaliteit doorstroom en dubbele finaliteit) en 4.6 (Finaliteit arbeidsmarkt) De leerlingen beoordelen kritisch wederzijdse invloeden uit tussen enerzijds het individu en de samenleving en anderzijds media, digitale infrastructuur en digitale toepassingen. (Evalueren).

- 4.8 (Finaliteit doorstroom en dubbele finaliteit) en 4.7 (Finaliteit arbeidsmarkt) De leerlingen passen regels van de digitale wereld kritisch toe. (Toepassen).

11.1.2 Eindtermen digitale competentie en mediawijsheid, en het Chatbot-project

Eindterm 4.1 (tweede en derde graad) - Zelfvertrouwen

Men moet geen bolleboos zijn in wiskunde en wetenschappen, noch moet men een kei zijn in programmeren om met *Natural Language Processing* aan de slag te gaan.

Door de leerlingen positieve ervaringen te bieden in een voor hen waarschijnlijk ongekend domein dat echter maatschappelijk relevant is en gelinkt is aan hun leefwereld - leerlingen maken bv. veel gebruik van sociale media -, kan hun zelfvertrouwen betreffende informaticawetenschappen, in het bijzonder programmeren en artificiële intelligentie, een boost krijgen.

De leerlingen uit niet-STEM-richtingen zien door te werken met de notebooks van het project 'Chatbot' in dat taaltechnologie ook voor hen een haalbaar domein is.

Ook leerlingen uit STEM-richtingen krijgen meer zelfvertrouwen doordat ze erin slagen gekende principes van programmeren toe te passen in een nieuwe probleemstelling.

Eindterm 4.2 (tweede en derde graad) - Compatibiliteit tussen digitale infrastructuur en toepassingen (tweede graad) - Standaardfunctionaliteiten/Functionaliteiten

In de notebooks van het Chatbot-project wordt er gewerkt met strings. Dat is het type dat Python gebruikt voor tekst. Die tekst wordt dan bewerkt en/of verwerkt door eenvoudige functies aan te wenden.

De leerlingen maken kennis met bestanden met de extensie 'ipynb', die enkel met geschikte software geopend kunnen worden.

Bij de sentimentanalyse moet de tekst in het juiste formaat worden aangeboden aan het AI-systeem, anders kan het AI-systeem niets doen. Bv. in de notebook (die gebruikmaakt van een regelgebaseerd AI-systeem) wordt de tekst voorverwerkt, zoals het vervangen van de hoofdletters door kleine letters en het terugbrengen van de woorden tot hun woordenboekvorm.

Men kan aan de leerlingen meegeven dat ook in andere *Natural Language Processing*-toepassingen de tekst waarmee een AI-systeem moet werken, in het juiste formaat aangeboden moet worden. Dat ligt voor de hand bij automatische cyberpestdetectie, aangezien dat eigenlijk volgens dezelfde principes werkt als sentimentanalyse. Maar

ook bij digitale assistenten is het formaat van de data belangrijk: een digitaal geluidsbestand moet eerst worden omgezet naar tekst d.m.v. spraakherkenningssoftware (spraak naar tekst) vooraleer de assistent de vraag kan analyseren en vervolgens een passend antwoord kan genereren.

Eindterm 4.3 (tweede en derde graad) - Digitale structuur en toepassingen om digitaal te communiceren

De chatbot is een voorbeeld van digitale communicatie. Door de chatbot-notebook te doorlopen, begrijpen de leerlingen de beperkingen en de mogelijkheden van zo'n systeem en zijn ze beter in staat een chatbot op een website op een adequate manier te gebruiken.

Het project 'Chatbot' geeft ook suggesties voor mogelijke activiteiten in de klas die dit inzicht bevorderen.

Door zich te verdiepen in de sentimentanalyse worden de leerlingen zich meer bewust van wat er allemaal komt kijken bij de sociale media en kunnen ze er op een meer bewuste manier berichten posten.

Eindterm 4.4 (tweede graad) - De bit als basiseenheid van data

Data worden aan de computer gegeven in een digitaal formaat. Concreet betekent dit dat de data worden omgezet naar getallen, die op hun beurt worden voorgesteld door een rij van nullen en enen. Zo'n nul of zo'n één is een bit. Bv. de natuurlijke getallen van 0 t.e.m. 255 kunnen worden voorgesteld door een rij van 8 nullen en enen, bv. 35 is 00100011 (een rij van acht nullen en enen is een byte).

Bijvoorbeeld, *one hot encoding*: bij tekst wordt aan de computer een lijst van woorden meegegeven in een bepaalde volgorde. Met een rij van nullen en een één, bijvoorbeeld 000000000000000000001000000000...0, kan een bepaald woord dan aan de computer gegeven worden. In het voorbeeld staat de 1 op de zestiende plaats in de rij wat betekent dat het woord het zestiende woord in de woordenlijst is. Om aan de computer kenbaar te maken welke van de woorden in de woordenlijst voorkomen in een bepaalde klantenreview, kan dan een rij nullen en enen, zoals 0010111010000001000000100...0, worden meegegeven. Dat betekent dat het derde, vijfde, zesde, zevende, negende, zestiende, drieëntwintigste ... woord in de review staat.

Eindterm 4.4 (tweede en derde graad) - Zenden, ontvangen, verwerken, opslaan, afspraken tussen digitale systemen

Om aan de slag te gaan met een notebook, wordt de URL ingegeven in de webbrowser. Deze URL bevat een protocol (HTTPS; wat wijst op een veilige verbinding) en een

domeinnaam (kiks.ilabt.imec.be/jupyter).

De webbrowser legt via het internet contact met een DNS-server om het IP-adres (een numeriek internetadres) op te zoeken van de server van AI Op School. Vervolgens wordt de boodschap van de gebruiker (de *client*, hier de webbrowser van de computer van de leerling of leerkracht) doorgegeven aan de webserver van AI Op School. Er wordt namelijk aan deze server gevraagd om toegang te geven tot de notebooks van AI Op School.

Het *Hypertext Transfer Protocol* (HTTP) is een computerprotocol waarin afspraken zijn vastgelegd zodat een computer (in dit geval de computer van de leerling of leerkracht) en een andere computer (hier de webserver van AI Op School) met elkaar kunnen communiceren.

HyperText Transfer Protocol Secure (HTTPS) is niets anders dan een uitbreiding van het HTTP-protocol om de uitwisseling van gegevens ook op een veilige manier te laten gebeuren. Het HTTPS-protocol versleutelt de gegevens.

De server geeft dan toegang tot de verschillende projecten die op de server staan. Het gewenste project kan worden geselecteerd en nadien ook de gewenste notebook. Bij het doorlopen van de notebook communiceert de webbrowser voortdurend met de server, bv. bij het uitvoeren van een Python-script. Hierbij maakt de notebook gebruik van de CPU (rekeneenheid) of de GPU (nog meer rekenkracht) van de server.

De leerling kan een notebook downloaden en opslaan op een eigen computer. Om de notebook opnieuw te kunnen openen op de eigen computer, zal op die computer wel geschikte software geïnstalleerd moeten zijn.

Als een leerling een notebook wil uitprinten, dan moet de computer van de leerling kunnen communiceren met de printer. Ook hier is een protocol voor nodig.

Eindterm 4.4 (tweede graad) - Algoritme

De Python-scripts in de notebooks bevatten algoritmes die werden opgesteld om bepaalde doelen te bereiken.

Voorbeelden:

- In een notebook over sentimentanalyse gaat een algoritme elk woord in een review één voor één af en zet het om naar zijn woordenboekvorm.
- In de chatbot-notebook checkt een algoritme of de gestelde vraag voorkomt in de dataset, waarop hij een overeenkomstig antwoord teruggeeft.

Er zijn ook enkele notebooks die voorbereiden op de notebooks over sentimentanalyse en over chatbots. In deze notebooks passen de leerlingen een gegeven algoritme aan of stellen ze er zelf een op.

Eindterm 4.4 (tweede graad) - Input, verwerking, output

Voorbeelden:

- In de notebooks van sentimentanalyse vertrekt men van een klantenreview (de input). Deze gegeven tekst wordt verwerkt: de tekst wordt voorverwerkt waarna hij in het juiste formaat kan worden aangeboden aan het AI-systeem. Vervolgens analyseert het AI-systeem de review. Dit leidt tot slot tot een output: de review is positief, negatief of neutraal.
- Bij de notebook over de chatbot geeft de gebruiker een vraag in; dat is de input. Vervolgens speurt het AI-systeem in de databank naar die vraag of een die er niet teveel van verschilt; dat is de verwerking. Tot slot geeft de chatbot een antwoord aan de gebruiker, dat antwoord is de output.

Eindterm 4.4 (derde graad) - Databanken - - Beoordelen bouwstenen

Door de chatbot-notebook te doorlopen en de activiteiten rond chatbots te doen, zien leerlingen in dat nog veel van de huidige chatbots werken a.d.h.v. een groot databestand met vragen en antwoorden erin.

In de context van chatbots kunnen de tekstgenererende AI-systemen zoals GPT-2 en GPT-3 worden besproken. Dankzij het internet en *crawlers* is men in staat immense databanken aan te leggen waarop deze systemen worden getraind.

Eindterm 4.4 (derde graad) - Analoge en digitale voorstellingen

Zoals eerder vermeld wordt bij digitale assistenten een digitaal geluidsbestand eerst omgezet naar tekst d.m.v. spraakherkenningssoftware (spraak naar tekst) vooraleer de assistent de vraag kan analyseren en vervolgens een passend antwoord kan genereren. Dat digitaal geluidsbestand is zelf al een omzetting van een opname van een bepaald stemgeluid (een analogo geluidsbestand). Door de omzetting van analogo naar digitaal gaat steeds informatie verloren.

Eindterm 4.5 (tweede en derde graad) - Concepten van computationeel denken

In de notebooks van sentimentanalyse: decompositie (verschillende zaken moeten worden aangepakt bij het voorbereiden van de data) - algoritme - abstractie (gebruik van functies).

In de chatbot-notebook: decompositie (vraag - antwoord; deelprobleem: wat als de gestelde vraag niet in de dataset zit).

In de inhoud van project 'Chatbot': patroonherkenning (sentimentanalyse en cyberpestdetectie; chatbot en digitale assistent; plagiaatdetectie en auteursherkenning;

profilering (sollicitatie - gepersonaliseerde reclame en nieuws)).

Eindterm 4.5 (derde graad - Finaliteit doorstroom en dubbele finaliteit) - Digitale representatie van informatie

Data en informatie zijn niet hetzelfde. Een computer werkt met data, die op de juiste manier gerepresenteerd moeten zijn. De informatie wordt door de mens uit de data afgeleid door de data te interpreteren.

Eindterm 4.5 (tweede en derde graad - Finaliteit doorstroom en dubbele finaliteit)- Debuggen

Leerlingen voeren de code voortdurend uit tijdens het doorlopen van de notebooks. Bij foutmeldingen dienen ze onmiddellijk de fout op te sporen en vervolgens de code aan te passen.

Bij de chatbot-notebook merken de leerlingen dat de chatbot niet steeds zo goed functioneert. Daar passen ze de code aan om de effectiviteit van het script te verbeteren.

Eindterm 4.5 (tweede en derde graad - Finaliteit doorstroom en dubbele finaliteit) - Principes van programmeren

Sequentie - herhalingsstructuren - keuzestructuren komen aan bod in de notebooks.

Eindterm 4.5 (tweede en derde graad - Finaliteit doorstroom en dubbele finaliteit) - Elementen van programmeertalen

Begrippen zoals types, variabelen, operatoren, functies komen aan bod in de notebooks.

Om de strings te bewerken wordt bv. de +-operator gebruikt. Iets op het scherm laten verschijnen, gebeurt met de functie print().

Eindterm 4.6/4.5 (tweede en derde graad) - Wederzijdse invloeden tussen individu en media en maatschappij op leren, werken en vrije tijd

In de inhoud van het Chatbot-project komen deze wederzijdse invloeden ruimschoots aan bod.

Bedrijven en organisaties screenen socialemediaposts om te proberen te achterhalen hoe tevreden hun klanten zijn van hun diensten en producten. Ze kunnen de informatie die ze bekomen door de automatische sentimentanalyse gebruiken om hun diensten en producten eventueel aan te passen. Ook politieke partijen en de overheid detecteren (on)tevredenheid over hun standpunten of hun beleid, en baseren zich hierop om

eventueel dat beleid bij te sturen of bepaalde intenties toch niet of juist wel door te voeren.

Posts op sociale media worden door het betreffende platform zorgvuldig bijgehouden en blijvend gelinkt aan de gebruiker. Aan de hand van het profiel van de gebruiker dat men op die manier opmaakt, worden gepersonaliseerde reclame en gepersonaliseerd nieuws aan de gebruiker aangeboden. Ook op basis van het koop- en klikgedrag van gebruikers met ‘verwante’ interesses wordt de ervaring op het socialemediaplatform gepersonaliseerd. Aanbevelingsalgoritmes kunnen zo bijvoorbeeld een sterke invloed uitoefenen op het online shopgedrag (e-commerce).

Omdat aanbevelingsalgoritmes ook steeds meer gebruikt worden op websites van online kranten en ook het resultaat van een zoekmachine steeds meer gepersonaliseerd wordt, is men best alert voor de filterbubbel. Als men onderhevig is aan een filterbubbel kan het beeld dat men heeft van de wereld misvormd zijn. Als men bv. wil leren over een onderwerp waar men niet vertrouwd mee is, dan kan men door zo'n bubbel een ongenueanceerd beeld krijgen.

Dankzij automatische vertaling kan leren gemakkelijker en rijker worden, bv. omdat teksten in een andere taal nu ook toegankelijk zijn.

Eindterm 4.6/4.5 (tweede en derde graad) - Datawijsheid

Leerlingen beseffen door de projectinhouden hoe data die zij achterlaten op het world wide web door verschillende partijen kunnen worden aangewend en met elkaar in verband kunnen worden gebracht. Ze moeten beseffen dat omdat de reclame gepersonaliseerd is, ze er mogelijk vatbaarder voor zijn.

Eindterm 4.7/4.6 (tweede en derde graad) - Ethiek

Leerlingen zijn zich ervan bewust dat met nieuwe technologieën ook ethische aspecten kunnen opduiken. In het Chatbot-project worden voorbeelden van ethische aspecten aangereikt die gebruikt kunnen worden als onderwerp voor een klasgesprek.

Eindterm 4.7/4.6 (tweede en derde graad) - Anonimiteit op het wereldwijde web

Anonimiteit op het web kan drempels wegnemen, bv. om te gaan cyberpesten of om haatspraak te verspreiden. Anderzijds is men m.b.v. AI-systemen soms wel al in staat de anonieme auteur van een bepaalde tekst te achterhalen.

Eindterm 4.7/4.6 (tweede en derde graad) - Dynamisch karakter van de digitale wereld

Voorbeeld: digitale assistenten geraken steeds meer ingeburgerd thuis, via smartphone,

in de auto. Robots komen steeds meer in beeld. Er is een wisselwerking tussen maatschappelijke noden en ontwikkelingen in de computerwetenschappen.

Eindterm 4.8/4.7 (tweede en derde graad) - Risico's mediagedrag

In het Chatbot-project worden voorbeelden gegeven van nieuwe praktijken die steeds meer hun weg vinden naar de sector van human resource, zoals het screenen van sociale media. Leerlingen beseffen dat wat op socialemediaplatformen gepost wordt, er binnen enkele jaren als zij gaan solliciteren nog steeds te vinden is.

11.2 STEM

11.2.1 STEM-eindtermen

Deze eindtermen behoren tot de eindtermen Wiskunde - natuurwetenschappen - technologie - STEM. Ze zijn op dezelfde manier geformuleerd voor de tweede en de derde graad. Dat geldt niet noodzakelijk voor de conceptuele en procedurele kennis.

- 6.54/6.40 (Finaliteit doorstroom) *De leerlingen analyseren natuurlijke en technische systemen aan de hand van verschillende STEM-concepten. (Analyseren).*
- 6.15/6.17 en 6.30/6.21 (resp. Finaliteit arbeidsmarkt en dubbele finaliteit) *De leerlingen analyseren natuurlijke en technische systemen aan de hand van aangereikte STEM-concepten. (Analyseren).*
- 6.55/6.41 en 6.31/6.22 (resp. Finaliteit doorstroom en dubbele finaliteit) *De leerlingen ontwerpen een oplossing voor een probleem door concepten en praktijken uit verschillende STEM-disciplines geïntegreerd aan te wenden. (Creëren).*
- 6.16/6.18 (Finaliteit arbeidsmarkt) *De leerlingen ontwerpen een oplossing voor een probleem door concepten en praktijken uit verschillende STEM-disciplines geïntegreerd aan te wenden. (Creëren).*
- 6.56/6.42 (Finaliteit doorstroom) *De leerlingen beargumenteren vanuit verschillende invalshoeken keuzes bij het ontwerp en het gebruik van technische systemen en andere STEM-oplossingen. (Evalueren).*
- 6.17/6.19 en 6.32/6.23 (resp. Finaliteit arbeidsmarkt en dubbele finaliteit) *De leerlingen beargumenteren vanuit verschillende invalshoeken keuzes bij het gebruik van technische systemen. (Evalueren).*
- 6.57/6.43 (Finaliteit doorstroom) *De leerlingen onderzoeken aan de hand van concrete maatschappelijke uitdagingen de wisselwerking tussen STEM-disciplines onderling en tussen STEM-disciplines met de maatschappij. (Analyseren).*

- *6.18/6.20 en 6.33/6.24 (resp. Finaliteit arbeidsmarkt en dubbele finaliteit) De leerlingen leggen aan de hand van concrete maatschappelijke uitdagingen de samenwerking tussen STEM-disciplines onderling en tussen STEM-disciplines met de maatschappij uit. (Begrijpen).*

11.2.2 STEM-eindtermen en het Chatbot-project

Eindtermen - STEM-concepten

Door de opkomst van het internet zijn grote stromen van gegevens op gang gekomen. Uit de big data die men verkrijgt op sociale media kan men veel informatie inwinnen (informatie). Het posten van een bericht op sociale media kan gevolgen hebben die men op voorhand niet heeft ingeschat (oorzaak - gevolg). De AI-systemen maken gebruik van patroonherkenning. A.d.h.v. AI-systemen worden profielen opgesteld van bv. sollicitanten, gebruikers van webshops en online krantenabonnees (patronen - systemen). Door het optreden van een filterbubbel kunnen bv. het verwerven van kennis en de democratie verstoord worden (stabiliteit en verandering). De sociale netwerken die reeds offline bestaan kennen een enorme schaalvergroting door de opkomst van de sociale media, met bv. het spontaan ontstaan van het domein van de sentimentanalyse. Door het inzetten van chatbots wordt bekeken hoe een klantendienst een betere dienstverlening kan verzekeren, en bovendien aan meer klanten, 24 uur op 24 uur (verhouding en hoeveelheden).

De leerlingen zien in dat de AI-systemen voor sentimentanalyse die gebruikt worden in de notebooks, beperkingen hebben. Ze kunnen bv. moeilijk om met ironie. Dat ligt natuurlijk aan het feit dat er met een lexicon wordt gewerkt waarin een bepaalde sentimentscore aan de woorden is toegekend zonder rekening te houden met de context. De manier waarop het AI-systeem ontworpen is, ligt dus aan de basis van zijn beperkingen.

Een ML-systeem gebruiken voor sentimentanalyse kan mogelijk deze beperkingen gedeeltelijk opvangen.

Aanbevelingsalgoritmes doen aanbevelingen die bv. gebaseerd zijn op bepaalde patronen die aanwezig zijn in het klik- en koopgedrag van de gebruiker.

AI-systemen om te monitoren wat er gepost wordt op socialemediaplatformen en in online winkels worden steeds meer aangewend. Ze geven bedrijven en organisaties informatie die kan leiden tot aanpassingen (veranderingen) in hun werking.

Chatbots zijn doorgaans nog regelgebaseerde systemen, die daardoor slechts zeer specifiek bruikbaar zijn. De leerlingen onderzoeken de effectiviteit van sommige

chatbots.

Om de chatbots ook vragen te laten herkennen die gelijken op wat aanwezig is in de dataset worden zogenaamde afstandsfuncties uit ML gebruikt. De 'afstand' tussen gelijkende vragen is dan klein. Hiervoor worden wiskundige concepten gebruikt. Maar GPT-2 en GPT-3 zouden in de toekomst wel een rol kunnen gaan spelen in de ontwikkeling van meer gesofisticeerde chatbots. Omdat deze systemen getraind zijn met data afkomstig vanop het world wide web, zijn ze extra gevoelig voor bias.

De leerlingen zien in dat er een wisselwerking is tussen de impact van het internet en de groeiende aanwezigheid van AI-systemen enerzijds, en de maatschappij, waaronder de economie en hoe met ethische aspecten wordt omgegaan, anderzijds.

In het Chatbot-project wordt bijvoorbeeld ook getoond hoe *Natural Language Processing* kan helpen bij historisch onderzoek, bv. het inzetten van automatische auteursherkenning bij Latijnse teksten.

11.3 Nederlands

11.3.1 Eindtermen Nederlands derde graad

- 2.5 De leerlingen selecteren relevante informatie in geschreven en gesproken teksten in functie van doelgerichte informatieverwerking en communicatie. (Analyseren).
- 2.12 (Finaliteit doorstroom en dubbele finaliteit) De leerlingen gebruiken het inzicht in de regels en kenmerken van het Standaardnederlands als taalsysteem ter ondersteuning van doelgerichte informatieverwerking en communicatie. (Toepassen).
- 2.11 (Finaliteit arbeidsmarkt) De leerlingen gebruiken de belangrijkste regels en kenmerken van het Standaardnederlands als taalsysteem ter ondersteuning van doelgerichte informatieverwerking en communicatie. (Toepassen).

11.3.2 Eindtermen Nederlands en het Chatbot-project

Eindterm 2.5 - Informatie selecteren

In de notebooks voor sentimentanalyse selecteren de leerlingen m.b.v. Python de afzonderlijke woorden in een review, om die vervolgens te matchen met een lexicon, en zo de sentimentwaarde van de review te bepalen.

Eindterm 2.12/2.11 - Samenstellingen en afleidingen, woordsoort, letterlijk en figuurlijk, beeldspraak, vormen van humor, gevoelswaarde van woorden

In de notebooks voor sentimentanalyse worden woorden opgezocht in een lexicon. Om dat te kunnen doen, moeten die woorden eerst teruggebracht worden naar hun woordenboekvorm. Hierbij wordt ook gekeken naar de woordsoort, omdat een woord naargelang zijn woordsoort een andere sentimentwaarde kan hebben.

De leerlingen merken dat AI-systemen voor sentimentanalyse bijvoorbeeld worstelen met ironie en sarcasme, en ze niet goed om kunnen gaan met beeldspraak.

Bij auteursherkenning zijn het vooral de kleine woordjes, zoals de lidwoorden en de voegwoorden, die de auteur aanwijzen.

Versie 1.0

Versie 1.0

BIBLIOGRAFIE

- Ainsworth, J. & Juola, P. (2019). Who Wrote This?: Modern Forensic Authorship Analysis as a Model for Valid Forensic Science. *Washington University Law Review*, 96. https://openscholarship.wustl.edu/law_lawreview/vol96/iss5/10.
- ALICE (1995). Tell a.i.i.c.e. Geraadpleegd op 7 april 2021 via <https://www.pandorabots.com/pandora/talk?botid=b8d616e35e36e881>.
- Bank of England (2021). *Biljet 50 pond met Alan Turing* [Afbeelding]. Geraadpleegd op 30 juli 2021 via <https://www.bankofengland.co.uk/news/2021/march/the-new-50-note-unveiled>.
- Bathina, K., ten Thij, M., Lorenzo-Luaces, L., Rutter, L. A., & Bollen, J. (2021). Individuals with depression express more distorted thinking on social media. *Nature Human Behaviour*.
- Berbers, Y., Debeuckelaere, W., Herdt, P. D., Desmedt, Y., Smet, F. D., Hildebrandt, M., Poels, K., Pierson, J., Preneel, B., & Vandewalle, J. (2017). *Privacy in tijden van internet, sociale netwerken en big data*. Brussel: KVAB Standpunt 49.
- Bergan, B. (2020). VR Reunites Mother with Deceased Daughter. Geraadpleegd op 2 augustus 2021 via <https://interestingengineering.com/vr-reunites-mother-with-deceased-daughter>.
- Berkers, E. (2020). Eliza. Geraadpleegd op 9 april 2021 via <https://www.electicenergies.com/nederlands/psyche/eliza>.
- Bingli (2021). Bingli, het slimme medische interview. Geraadpleegd op 1 augustus 2021 via <https://www.mybingli.com>.
- Boden, M. A. (2016). *AI. Its nature and future*. Oxford, Verenigd Koninkrijk: Oxford University Press.

- Brants, M. (2020). Vlaams Chatbot Onderzoek bij eindgebruiker. *Thomas More*. Geraadpleegd op 1 augustus 2021 via <https://www.chatbotgids.com/post/vlaams-chatbot-onderzoek-bij-eindgebruiker>.
- Byron, G. G. (2015). She walks in beauty. Geraadpleegd op 21 september 2019 via <https://www.poetryfoundation.org/poems/43844/she-walks-in-beauty>.
- Capgemini Research Institute (2019). Voice on the go. How can auto manufacturers provide a superior in-car voice experience. Geraadpleegd op 9 augustus 2021 via <https://www.capgemini.com/wp-content/uploads/2019/11/Report-%E2%80%93-Voice-on-the-Go.pdf>.
- Carl H. Pforzheimer Collection of Shelley and His Circle. The New York Public Library (1835). *Watercolor portrait of Ada Byron* [Afbeelding]. Geraadpleegd op 22 september 2019 via <http://digitalcollections.nypl.org/items/8562182e-4845-db2b-e040-e00a18060ea6>.
- Centraal Bureau voor de Statistiek (2020). Corona sentimentsindicator. Geraadpleegd op 10 augustus 2021 via <https://www.cbs.nl/nl-nl/over-ons/innovatie/project/corona-sentimentsindicator>.
- Chaski, C. E. (2005). Who's At The Keyboard? Authorship Attribution in Digital Evidence Investigations. *International Journal of Digital Evidence*, 4(1).
- Chin, C. & Robinson, M. (2020). How AI bots and voice assistants reinforce gender bias. Geraadpleegd op 9 augustus 2021 via <https://www.brookings.edu/research/how-ai-bots-and-voice-assistants-reinforce-gender-bias/>.
- Chollet, F. (2018). *Deep learning with Python*. Shelter Island, NY: Manning Publications Co.
- Coeckelbergh, M. (2020). *AI ETHICS*. Cambridge, MA: The MIT Press.
- Criado Perez, C. (2019). *Onzichtbare vrouwen*. Prometheus.
- Cypher789 (2005). *Spam 2* [Afbeelding]. CC BY-SA 3.0. Geraadpleegd op 30 juli 2021 via https://commons.wikimedia.org/wiki/File:Spam_2.jpg.
- Davies, S. (2021). Love bytes. A chatbot provides emotional support to lonely hearts and potentially mines data from millions of vulnerable users. Geraadpleegd op 9 april 2021 via <https://www.theworldofchinese.com/2021/03/love-bytes/>.

- De Clercq, O. & Van Hee, C. (2021). Sentiment Demo. Geraadpleegd op 5 april 2021 via <https://lt3.ugent.be/resources/sentiment-demo/>.
- De Croo, A. (2021). [@alexanderdecroo]. *Wat gisteren gebeurd is in het Terkamerenbos is totaal onaanvaardbaar. Alle steun aan gekwetste politiemensen. Ik begrijp dat iedereen corona beu is. Maar de regels zijn er met reden en voor iedereen. Ziekenhuizen lopen vol. Solidariteit nu is sleutel voor vrijheid morgen.* [Tweet]. Twitter. <https://twitter.com/alexanderdecroo/status/1377918008761454592>.
- De Gussem, J. (2019). Collaborative authorship in twelfth-century Latin literature: a stylometric approach to gender, synergy and authority.
- De Gussem, J. (2020a). Hoe taalalgoritmes een middeleeuwse liefdesaffaire bekoelden. Geraadpleegd op 31 maart 2021 via <https://www.eoswetenschap.eu/geschiedenis/hoe-taalalgoritmes-een-middeleeuwse-liefdesaffaire-bekoelden>.
- De Gussem, J. (2020b). *Hoe taalalgoritmes een middeleeuwse liefdesaffaire bekoelden* [Videobestand]. Geraadpleegd op 30 maart 2021 via <https://www.youtube.com/watch?v=gI7a0yn1mw8>.
- De Gussem, J. & Daelemans, W. (2020). Ben je wat je schrijft? *EOS Psyche & Brein Special*, (pp. 40–43).
- De Ketelaere, G. M. (2020). *Mens versus machine*. Pelckmans.
- De Winne, I. (2020). *Taalgebruik op CV onthult persoonlijkheid* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://kanaalz.knack.be/nieuws/taalgebruik-op-cv-onthult-persoonlijkheid/video-normal-1678309.html>.
- Dehaene, T. (2020). Dutch GPT2: Autoregressive Language Modelling on a budget. Geraadpleegd op 12 juni 2021 via <https://blog.ml6.eu/dutch-gpt2-autoregressive-language-modelling-on-a-budget-cff3942dd020>.
- Dickson, B. (2019). What is Natural Language Processing (NLP). Geraadpleegd op 31 maart 2021 via <https://www.pcmag.com/opinions/what-is-natural-language-processing-nlp>.
- Eckert, M. (2019). Nip-test screent ook moeder op genetische defecten. *De Standaard*. Geraadpleegd via Gopress op 9 september 2019 via https://www.standaard.be/cnt/dmf20190904_04592057.

- ELIZA (2018). *ELIZA conversation* [Afbeelding]. Publiek domein. Geraadpleegd op 28 juli 2019 via https://commons.wikimedia.org/wiki/File:ELIZA_conversation.jpg.
- Emmery, C., Verhoeven, B., De Pauw, G., Jacobs, G., Van Hee, C., Lefever, E., Desmet, B., Hoste, V., & Daelemans, W. (2020). Current limitations in cyberbullying detection: On evaluation criteria, reproducibility, and data scarcity. *LANGUAGE RESOURCES AND EVALUATION*, (pp.37).
- Enix, S. (2018). *Shadow-of-the-Tomb-Raider-010419-002* [Afbeelding]. Geraadpleegd op 6 september 2019 via <https://www.flickr.com/photos/161227653@N02/47513423051/>. Publiek domein.
- Eynikel, J. (2017). *Robot aan het stuur. Over de ethiek van techniek*. Tielt: Lannoo.
- FBI (2021). Famous Cases and Criminals. The Unabomber. Geraadpleegd op 5 april 2021 via <https://www.fbi.gov/history/famous-cases/unabomber>.
- Feine, J., Morana, S., & Gnewuch, U. (2019). Measuring Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis. *14th International Conference on Wirtschaftsinformatik*.
- Furhat Robotics (2018). *Unbiased Recruiter Robot - Furhat Robotics* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://youtu.be/rPKrdxiEkQ0>.
- Futureism (2017). *This 'Racist soap dispenser' at Facebook office does not work for black people* [Videobestand]. Geraadpleegd op 8 augustus 2020 via https://youtu.be/YJjv_0eiHmo.
- Følstad, A. & Brandtzaeg, P. B. (2020). Users' experiences with chatbots: findings from a questionnaire study. *Qual User Exp*, 5(3).
- Gabriels, K. (2016). *Onlife*. Lannoo.
- Gabriels, K. (2019). *Regels voor robots. Ethiek in tijden van AI*. Academic And Scientific Publishers.
- Garderen, F. V. (2019). Nog steeds klachten over gebrekkige begeleiding: veel zwangere vrouwen weten niet wat NIPT inhoudt. *De Morgen*. Geraadpleegd 15 september 2019 via <https://www.demorgen.be/nieuws>.
- Google (2020). *Embeddings can produce remarkable analogies* [Afbeelding]. CC 4.0. Geraadpleegd op 21 augustus 2021 via <https://developers.google.com/>

machine-learning/crash-course/embeddings/translating-to-a-lower-dimensional-space.

- GPT-3 (2020). A robot wrote this entire article. Are you scared yet, human? Geraadpleegd op 6 augustus 2021 via <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>.
- Grossman, D. (2020). VR reunites a woman with her dead daughter. VR therapy taken to a whole new level of customization. Geraadpleegd op 2 augustus 2021 via <https://www.inverse.com/innovation/korean-vr-reunites-family-beyond-grave>.
- Güzeldere, G. & Franchi, S. (1995). Dialogues with colorful “personalities” of early AI. *Stanford Hum. Rev.*, 4(2), 161–169.
- Health, W. (2021). Welcome to the future of mental health. Geraadpleegd op 9 april 2021 via <https://woebothealth.com/>.
- IDLab (2021). Pretturn. Geraadpleegd op 5 april 2021 via <http://predict.idlab.ugent.be/projects/Pretturn/>.
- IntelligenceSquared Debates (2019). *IBM Project Debater* [Videobestand]. Geraadpleegd op 4 april 2021 via https://youtu.be/3_yy0dnIc58.
- Jia, Y. & Weiss, R. (2019). Introducing Translatotron: An End-to-End Speech-to-Speech Translation Model. Geraadpleegd op 8 augustus 2021 via <https://ai.googleblog.com/2019/05/introducing-translatotron-end-to-end.html>.
- Joris, G. (2021). Politiek nieuws, artificiële intelligentie en het belang van diversiteit. In N. Bouteica & J. De Bruyne (Eds.), *Artificiële intelligentie en maatschappij* (pp. 115–132). Gompel & Svacina.
- Juola, P. (2012). Detecting Stylistic Deception. *Proceedings of the EACL 2012 Workshop on Computational Approaches to Deception Detection*, (pp. 91–96). Association for Computational Linguistics.
- Juola, P. (2013). How Computer Program Helped Show J.K.Rowling write A Cuckoo's Calling: Author of the Harry Potter books has a distinct linguistic signature. *Scientific American*, 20.
- Juola, P. (2015). The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions. *Digital Scholarship in the Humanities*, 30, i100–i113.

- Juola, P. (2017). DETECTING CONTRACT CHEATING VIA STYLOMETRIC METHODS. *Conference Proceedings on Plagiarism across Europe and Beyond*, (pp. 187–198).
- Kenniscentrum Data & Maatschappij (2020a). Chatbots meer dan trouwe medewerkers van een helpdesk. Geraadpleegd op 31 maart 2021 via <https://data-en-maatschappij.ai/nieuws/chatbots-meer-dan-trouwe-medewerkers-van-een-helpdesk>.
- Kenniscentrum Data & Maatschappij (2020b). De chatbot van het Kenniscentrum Data & Maatschappij. Geraadpleegd op 31 maart 2021 via <https://data-en-maatschappij.ai/nieuws/de-chatbot-van-het-kenniscentrum-data-maatschappij>.
- Kenniscentrum Data & Maatschappij (2020c). Tien tips om zeker te zijn dat je chatbot een meerwaarde heeft voor je doelgroep. Geraadpleegd op 31 maart 2021 via <https://data-en-maatschappij.ai/nieuws/tien-tips-om-zeker-te-zijn-dat-je-chatbot-een-meerwaarde-heeft-voor-je-doelgroep>. Frederic Heymans.
- Kenniscentrum Data & Maatschappij (2021a). *brAInfood Chatbots* [Afbeelding]. Geraadpleegd op 31 maart 2021 via <https://data-en-maatschappij.ai/publicaties/brainfood-databescherming-en-chatbot>.
- Kenniscentrum Data & Maatschappij (2021b). *brAInfood Voice assistants*. Geraadpleegd op 31 maart 2021 via <https://data-en-maatschappij.ai/publicaties/brainfood-databescherming-en-voice-assistants-1>.
- Kestemont, M., Stover, J., Koppel, M., Karsdorp, F., & Daelemans, W. (2016). Authenticating the writings of Julius Caesar. *Expert Systems with Applications*, 63, 86–96.
- Kim, V. (2020). Virtual Reality, Real Grief. Geraadpleegd op 2 augustus 2021 via <https://slate.com/technology/2020/05/meeting-you-virtual-reality-documentary-mbc.html>.
- Koh, D. (2019). Ping An Good Doctor launches commercial operation of One-minute Clinics in China. Geraadpleegd op 31 maart 2021 via <https://www.mobihealthnews.com/news/apac/ping-good-doctor-launches-commercial-operation-one-minute-clinics-china>.
- Kröger, F. (2016). *Autonomous Driving. Automated Driving in Its Social, Historical and Cultural Contexts*. Berlin, Heidelberg: Springer.

- Legg, S. & Hutter, M. (2007). Universal Intelligence: A Definition of Machine Intelligence. *Minds & Machines*, 17(4), 391–444.
- Lieman, R. (2019). Gepersonaliseerd nieuws: matchmaker voor online media of journalistiek-ethisch mijnenveld? Geraadpleegd op 7 april 2021 via <https://www.vn.nl/gepersonaliseerd-nieuws-matchmaker-of-mijnenveld/>.
- Lim, M. (2018). History of AI Winters. Geraadpleegd op 27 december 2018 via <https://www.actuaries.digital/2018/09/05/history-of-ai-winters/>.
- Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.
- Luka, Inc (2020). Replika: My AI Friend. Geraadpleegd op 9 april 2021 via <https://play.google.com/store/apps/details?id=ai.replika.app>.
- Lyu, J., Le Han, E., & Luli, G. (2021). COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis. *Journal of Medical Internet Research*, 23.
- Microsoft (2019). *Demo: The magic of AI neural TTS and holograms at Microsoft Inspire 2019* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://youtu.be/auJJrHgG9Mc>.
- Nardone, C. (2007). *Shakey the Robot* [Afbeelding]. CC BY-SA 2.0. Geraadpleegd op 28 juli 2019 via [https://commons.wikimedia.org/wiki/File:Shakey_the_Robot_\(developed_between_1966-1972_at_SRI_International\)_-_Computer_History_Museum_\(2007-11-10_23.16.01_by_Carlo_Nardone\).jpg](https://commons.wikimedia.org/wiki/File:Shakey_the_Robot_(developed_between_1966-1972_at_SRI_International)_-_Computer_History_Museum_(2007-11-10_23.16.01_by_Carlo_Nardone).jpg).
- Neff, G. & Nagy, P. (2016). Talking to Bots: Symbiotic Agency and the Case of Tay. *International Journal of Communication*, 10, 4915–4931.
- Ng, K. (2020). 'Maybe it's a real paradise': Mother reunited with dead seven-year-old daughter in virtual reality. Geraadpleegd op 2 augustus 2021 via <https://www.independent.co.uk/news/world/asia/dead-child-virtual-reality-mother-daughter-south-korea-jang-ji-sung-a9325691.html>.
- NJIT IST (2018). ELIZA: a very basic Rogerian psychotherapist chatbot. Geraadpleegd op 8 april 2021 via <https://web.njit.edu/~ronkowit/eliza.html>.
- Oostra, L. (2019). AI bepaalt of trainees bij Unilever aangenomen worden: 'Je cv werkt niet meer voor je'. Geraadpleegd op 6 september 2019 via <https://www.mt.nl/management/ai-trainees-unilever-cv-werkt-niet-meer-voor-je/570440>.

- OpenAI (2019). Better Language Models and Their Implications. Geraadpleegd op 31 maart 2021 via <https://openai.com/blog/better-language-models>.
- Pettersson, M. (2021). [@mariapettersson]. *Kyllä, algoritmi voi olla seksisti, ja onkin. Se johtuu siitä, että me ihmiset olemme luoneet sen. Me siirrämme omat ennakoluulomme tekoälyyn ja koneisiin. Tässä esimerkkinä Google Translate. Koikeile itse!* [Tweet]. Twitter. <https://twitter.com/mariapettersson/status/1369022884878893059>.
- Ping An Good Doctor (2019). [@pagoooddoctor]. *Ping An Good Doctor meets the medical needs of drivers & passengers with the launch of OneminuteClinic at Nanjing Siqiao Expressway Service Area. The clinic will help improve road safety by providing drugs to sick people during their journey.* [Tweet] [Afbeelding]. Twitter. <https://twitter.com/PAGoodDoctor/status/1090097030926524418/photo/1>.
- Puschmann, C. & Powell, A. (2018). Turning Words Into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media. *Social Media + Society*, 4, 205630511879772.
- PwC (2018). *Benchmark studie over artificiële intelligentie, uitgevoerd in opdracht van het Departement Economie, Wetenschap en Innovatie van de Vlaamse Overheid. Rapport PwC.* Geraadpleegd op 29 mei 2019 via https://www.ewi-vlaanderen.be/sites/default/files/bestanden/departement_economie_wetenschap_en_innovatie_-_benchmark_studie_over_artificiele_intelligentie.pdf.
- Sakunkoo, P. & Sakunkoo, N. (2009). Analysis of Social Influence in Online Book Reviews. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1), 308–310.
- Schwartz, E. H. (2021). Google's Translatotron 2 Improves Linguistic Shifts Without the Deepfake Potential. Geraadpleegd op 8 augustus 2021 via <https://voicebot.ai/2021/07/27/googles-translatotron-2-improves-linguistic-shifts-without-the-deepfake-potential/>.
- Schwartz, H., Eichstaedt, J., Kern, M., Dziurzynski, L., Ramones, S., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M., & Ungar, L. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE*, 8(9).

- Science Museum Group Collection (1833). Portion of Babbage's calculating machine, Difference Engine no.1. Geraadpleegd op 22 september 2019 via <https://collection.sciencemuseumgroup.org.uk/objects/co62243/difference-engine-no-1-difference-engine-portion-only>.
- Shatz, E. (2018). *Alexa after man goes to dentist* [Videobestand]. Geraadpleegd op 30 maart 2021 via <https://www.youtube.com/watch?v=IKZToY-V16w>.
- Signorini, A., Segre, A., & Polgreen, P. (2011). The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. *PLOS ONE*, 6.
- Stad Gent (2021). Beleidsnota communicatie en merkbeleid 2020-2025. Geraadpleegd op 30 juli 2021 via <https://stad.gent/sites/default/files/media/documents/Beleidsnota%20Communicatie%20en%20merkbeleid%202020-2025.pdf>.
- Steels, L., Berendt, B., Pizurica, A., & Vandewalle, J. (2017). *Artificiële intelligentie. Naar een vierde industriële revolutie?* Brussel: KVAB Standpunt 53.
- Tatman, R. (2020). *Sociolinguistic Variation and Automatic Speech Recognition: Challenges and Approach* [Presentation]. Geraadpleegd op 8 april 2021 via http://www.rctatman.com/files/Tatman_2020_AAASocioAndASR.pd.
- Tazelaar, C. (1998). *Abélard. De briefwisseling met Héloïse*. pp. 20-21, 60-61. Ambo.
- Tech Insider (2018). *We Interviewed Pepper — The Humanoid Robot* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://www.youtube.com/watch?v=zJHyaD1psMc>.
- Tegmark, M. (2017). *LIFE 3.0*. New York, Verenigde Staten: Vintage Books, Penguin Random House.
- The Royal Society (2018). *What is artificial intelligence?* [Videobestand]. Geraadpleegd op 30 maart 2021 via <https://youtu.be/nASDYRkbQIY>.
- Thereforefilms (2016). *Sunspring*. Geraadpleegd op 8 april 2021 via <https://www.thereforefilms.com/sunspring.html>.
- Troyes, Médiathèque (13de eeuw, omgekende oorsprong). Recueil factice. MS 802. Geraadpleegd op 21 augustus 2021 via https://portail.mediathèque.grand-troyes.fr/iguana/www.main.cls?surl=search&p=*&recordId=2.2646&srchDb=2.

- van den Oord, A. (2017). High-fidelity speech synthesis with WaveNet [Blogpost]. Geraadpleegd op 31 maart 2021 via <https://deepmind.com/blog/article/high-fidelity-speech-synthesis-wavenet>.
- Van Hee, C., Jacobs, G., Emmerly, C., Desmet, B., Lefever, E., Verhoeven, B., Pauw, G., Daelemans, W., & Hoste, V. (2018). Automatic detection of cyberbullying in social media text. *PLOS ONE*, 13, e0203794.
- Van Royen, K., Poels, K., Daelemans, W., & Vandebosch, H. (2015). Automatic monitoring of cyberbullying on social networking sites: From technological feasibility to desirability. *Telematics and Informatics*, 32(1), 89 – 97.
- Vandekerckhove, S. (2019). *Nieuwe app voorspelt of u te vroeg zal bevallen. Gentse onderzoekers werken aan software die vroeggeborenen kan signaleren, maar ook kan helpen bij migraine of dementie*. Geraadpleegd op 5 april 2021 via GoPress. 7 augustus 2019.
- Vargha, D. (2021). [@doravargha]. *Hungarian is a gender neutral language, it has no gendered pronouns, so Google Translate automatically chooses the gender for you. Here is how everyday sexism is consistently encoded in 2021. Fuck you, Google*. [Tweet]. Twitter. <https://twitter.com/DoraVargha/status/1373211762108076034>.
- VLAIO (2020). Belgische primeur: AI-programma schrijft artikel voor IT-vakblad. Geraadpleegd op 4 augustus 2021 via <https://www.digitaletoeekomst.be/nl/artificiele-intelligentie/nieuws/belgische-primeur-ai-programma-schrijft-artikel-voor-it-vakblad>.
- VRT (2020). *Zo werkt liveondertiteling op de VRT* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://youtu.be/jhd7yxcA7Vw>.
- Walch, K. (2020). The Unique Challenges Of Voice Assistants In Vehicles. Geraadpleegd op 9 augustus 2021 via <https://www.forbes.com/sites/cognitiveworld/2020/09/29/the-unique-challenges-of-voice-assistants-in-vehicles/>.
- Wall Street Journal (2018). *Moving upstream. Artificial Intelligence: The Robots Are Now Hiring* [Videobestand]. Geraadpleegd op 4 april 2021 via <https://youtu.be/8QEK7B9GUhM>.
- Wikipedia (2019). Shakey the robot. Geraadpleegd op 26 juli 2019 via https://en.wikipedia.org/wiki/Shakey_the_robot.

- Williams, R., Machado, C., Druga, S., Breazeal, C., & Maes, P. (2018). "My doll says it's Ok": A study of Children's Conformity to a Talking Doll. *IDC '18. Proceedings of the 17th ACM Conference on Interaction Design and Children*, (pp. 625–631).
- Wolfram, S. (2015). Untangling the Tale of Ada Lovelace [Blog post]. Geraadpleegd op 21 september 2019 via <https://blog.stephenwolfram.com/2015/12/untangling-the-tale-of-ada-lovelace/>.
- ZombieTron (2021). [@zombietron]. *OK I did my own experiment and wrote "She is an engineer. He is beautiful" then swapped the translation back to English. That's mad sexist @Google* [Tweet]. Twitter. <https://twitter.com/ZombieTron/status/1373342849220100104>.

Versie 1.0

INDEX

- aanbeveling, 66, 119
aanbevelingsalgoritme, 100
Ada Lovelace, 25, 27
afstand, 139
AI-winter, 31, 32, 34
AI-zomer, 34
Alan Turing, 25
ALICE, 53
anamnesegeprek, 53, 54
annoteren, 22, 73, 83
artificiële intelligentie, 13, 15, 17, 35
autonome wagen, 111, 119
- bedrijf, 51, 56
belang, 74
bias, 103, 104, 117, 118, 121, 122
big data, 32
binaire classificatie, 83, 85
Bingli, 54
blog, 45, 47, 63, 80
broeikasgas, 114
- Caesar, 88, 89
chatbot, 51–55, 57, 60, 134
classificatie, 18, 48, 83, 85, 118
commercieel, 76
cyberpestdetectie, 131
cyberpesten, 81, 82, 86
- datagebaseerd, 16–18, 24, 60, 66, 68, 71, 73, 78, 80, 83, 86
deep learning, 18, 36, 40, 57, 89, 118
Deepl, 41
digitale assistent, 107, 108, 112, 132
diversiteitsalgoritme, 100, 105
dokter, 53, 54
domeinspecifiek, 71
dt-regels, 35
e-mail, 37
efficiënt, 56
elektronisch patiëntendossier, 38
ELIZA, 51, 52, 57
emotiedetectie, 58, 78
ethiek, 119
fake news, 114
film, 40
filter bubble, 100, 105
formaat, 32, 132
fysieke robot, 107, 112
gedicht, 40
geneeskunde, 76
general AI, 16
generaliseren, 19
genereren, 47
geschiedenis AI, 25
ghost writing, 97
Google Translate, 41
GPT, 42, 43, 45, 47, 48
Héloïse en Abélard, 90
hologram, 112
human resource, 103, 118
humor, 57
hybride, 69, 71
improvisatietoneel, 42, 108, 109
informatie, 36, 63
informatiebubbel, 99, 105
Instagram, 104
intakegesprek, 52
intelligentie, 13, 14
interview, 55
ironie, 57, 58, 71
juridisch, 39
kennisgebaseerd, 16, 17, 24, 57, 60, 66, 68, 70, 71, 80, 83, 86
klantendienst, 52
klantenreview, 63, 66, 76, 78, 80
klantentevredenheid, 55
kunstmatige intelligentie, 13, 15, 17, 35
label, 22, 73, 83, 100
labelen, 22
leeftijd, 104
lemma, 68, 70
lemmatisering, 68, 70
lerend algoritme, 18

- lexicon, 66, 68, 70, 71, 78, 80, 83, 86
- lowercasing, 67, 70
- machinaal leren, 17, 18, 24
- machine learning, 17, 18, 24, 64, 71, 73, 97, 108, 118
- matrix, 23
- medicatie, 38
- minen, 63
- module, 126
- monitoren, 75, 77, 82
- moreel, 119
- narrow AI, 16
- Natural Language Processing, 35, 57, 64, 82, 86, 99
- natuurlijke taal, 35, 51
- neurale netwerken, 89
- nieuwe technologie, 113, 115
- nieuws, 40
- nieuwsalgoritmes, 99
- NIPT, 55
- noden, 56
- notebook, 125, 126
- objectgeoriënteerd, 126
- ondertiteling, 40
- onderwijs, 54
- onderzoek, 55
- one hot encoding, 132
- one-minute-clinic, 54
- opinie, 64, 66, 75, 80
- opinion mining, 63
- organisatie, 51, 56, 76
- overfitting, 21
- part-of-speech, 68, 70
- participatie, 75
- peer review, 75
- personalisatie, 98, 99
- persoonlijkheid, 56, 66, 98, 103
- persoonskenmerken, 66
- Pinterest, 39
- plagiaat, 97
- podcast, 40, 41
- polariteit, 66, 68–71, 73, 78, 80
- politiek, 66, 77, 80
- preprocessing, 67, 71
- principal component analysis, 94, 97
- privacy, 65, 110, 116
- profilering, 98
- protocol, 133
- raster, 23
- realistisch stemgeluid, 109
- reclame, 65, 76, 98, 103
- regelgebaseerd, 57, 60, 70, 71, 78
- regressie, 18
- reinforcement learning, 21, 22
- rekenkracht, 31, 33, 114
- representatie, 23
- Rowling, 88
- sarcasme, 57
- schrijfstijl, 87, 105
- sentimentanalyse, 58, 63, 64, 74, 80, 83, 86, 107, 112, 131, 134
- sentimentscore, 69–71, 73, 75, 78
- server, 133
- singulariteit, 33
- slim speelgoed, 110
- sociale media, 63, 66, 75, 76, 81, 99, 103
- sociale netwerken, 64
- sollicitatie, 103, 118
- soorten AI, 16, 24
- spam, 37
- spraak naar spraak, 107, 109
- spraakherkenning, 108–111
- spraaksynthese, 111
- spraaktechnologie, 112
- statistiek, 17, 18, 87, 89
- supervised learning, 21
- Tay, 61
- technologie, 56
- tekstbegrip, 36
- tensor, 23
- terugkerende gebruiker, 56
- token, 67, 68, 70
- tokenisering, 67, 70
- transparantie, 119
- Turingtest, 25
- tweet, 63, 80
- Unilever, 103
- unsupervised learning, 21, 22
- valkuil, 71, 78
- vals negatief, 71
- vals positief, 71
- verkiezing, 66
- vertaling, 41
- voice-to-text, 108
- voorbereiden, 67, 131
- voorspellen, 18
- voorverwerken, 71
- vroeggeboorte, 38
- webbrowser, 133
- wereldkennis, 57
- wetenschappelijk onderzoek, 55, 64
- Woebot, 53, 116
- woordsoort, 68, 70
- word embedding, 37
- word2vec, 37
- Xiaoice, 54
- zekerheid, 18, 83
- zoekmachine, 37

Waar taal en technologie samenkomen, ontstaat het domein van Natural Language Processing (NLP). Met technieken uit machine learning, zoals het trainen van modellen om tekst, afbeeldingen of geluidsfrequenties te herkennen, kunnen computers op een intelligente manier werken met taal. Bekende NLP-toepassingen zijn bijvoorbeeld vertaalprogramma's, spraakbesturingssystemen, virtuele assistenten, en chatbots.

Hoe kan een AI-systeem teksten begrijpen, vertalen en zelfs schrijven? Kan een robot emoties of ironie herkennen in een tekst? Hoe belangrijk zijn data? Leerlingen van de tweede en de derde graad secundair onderwijs krijgen inzicht in het domein NLP en zetten hun eerste stappen in de wereld van het programmeren met Python.



9 789464 441000

