

Leren van andermans data

Citation for published version (APA):

Dekker, A. L. A. J. (2016). *Leren van andermans data*. Maastricht University.
<https://doi.org/10.26481/spe.20161117ad>

Document status and date:

Published: 17/11/2016

DOI:

[10.26481/spe.20161117ad](https://doi.org/10.26481/spe.20161117ad)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Unspecified

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Prof. dr. ir. Andre L.A.J. Dekker

Faculty of Health, Medicine and
Life Sciences

Leren van andermans data

Geachte mevrouw de rector, hooggeleerden, collega's, vrienden en familie. Bedankt dat u hier bent op mijn inaugurele rede als hoogleraar Clinical Data Science, klinische datawetenschappen.

Weet u, als je als hoogleraar wordt benoemd krijg je (naast een stropdas) een vademecum, een handleiding. Daarin staat onder andere wat over deze rede, de oratie. Een oratie is een openbare les, voor een breed publiek. Ik zal u de komende 45 minuten dus lesgeven. Maar het gebruik van visuele middelen zoals powerpoint in oraties wordt door de universiteit niet gewaardeerd. De jongste universiteit is niet zo van de moderne middelen. Het gaat om de redenaarskunst zegt het vademecum.

Om u toch enigszins houvast te geven heb ik de 45 minuten in 4 stukken gedeeld. Ik zal beginnen met een introductie en het waarom. Waarom wil ik leren van andermans data?

In het tweede stuk zal ik u wat vertellen over wat andermans data dan wel is en hoe je zorgt dat je ervan kan leren.

Dan over het leren zelf. Hoe doe je dat eigenlijk, leren van andermans data?

Het vierde en laatste deel gaat over wat je kan leren van andermans data en welke effecten dat hopelijk heeft.

Normaal sluit je dan af met een dankwoord. Maar deze universiteit is niet zo van de uitgebreide dankwoorden, ook dat staat in dat vademecum. Ik zal het dankwoord dan ook kort en algemeen houden, wees alstublieft niet teleurgesteld.

Leren van andermans data – het waarom

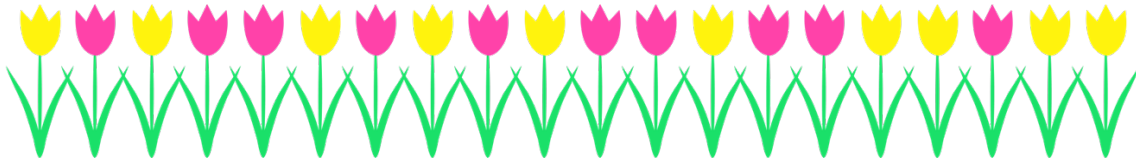
Laat ik niet beginnen met data maar met tulpen. Nou weet ik dat tulpen uit Holland en Maastricht niets met elkaar te maken hebben maar buiten onze landsgrenzen zijn dat soort gevoeligheden niet bekend en daar gebruik ik dit voorbeeld ook vaak. Stel je bent een tulpenkweker en je hebt een kratten gele en rode tulpenbollen gekocht. Maar de vakantiewerker die je hebt ingehuurd laat de kratten vallen en nu liggen die tulpenbollen allemaal door elkaar. Als tulpenkweker wil je dat er links op je veld rode tulpen staan en rechts gele zodat je ze goed kan oogsten en in rode en gele boeketten kan verkopen aan de Albert Heijn of in deze streek de Jan Linders. Je wil niet dat alle rode en gele dwars door elkaar staan want dat kun je niet goed oogsten.

Nu die vakantiewerker de boel door elkaar heeft gegooid, heb je een probleem want aan de buitenkant van een tulpenbol in het najaar kun je niet makkelijk zien welke kleur die wordt in het voorjaar. Dus huur je een expert in, een tulpenexpert, en je vraagt aan die expert: "Beoordeel elke bol en plant diegene, waarvan je het meest zeker bent dat die rood zal zijn, aan de linker kant. Plant diegene die zeker geel worden aan de rechterkant en plant diegene waarover je meest twijfelt in het midden.

Als je dan een perfecte tulpenexpert bent, ziet je veld er zo uit in april.



Dat is een perfecte voorspelling, 100% correct, in ons vakgebied noemen we dat een AUC van 1. Als je zoals als ik niets van tulpen weet, dan ziet het er zo uit, de helft van de tulpen staan fout, 50% correct, een AUC van 0.5.



En dit is een AUC van 0.7, er staan meer gele tulpen rechts dan links, maar het is niet perfect.



Dit noemen we een “classification task” experiment waarbij je twee groepen, klassen, van elkaar probeert te scheiden (in dit geval rode en gele tulpen) en dit voorbeeld gebruik ik om gevoel te geven voor die grootheid AUC. Hij geeft aan hoe goed een mens of een computer is in zo’n classificatie taak.

In Maastricht verbouwen we helemaal geen tulpen. Maar we behandelen bij MAASTRO Clinic waar ik werk, wel kankerpatiënten waaronder longkankerpatiënten. MAASTRO Clinic haar specialiteit is radiotherapie, het bestralen van kankerpatiënten, samen met chirurgie en chemotherapie de meest voorkomende behandeling voor kanker.

En Cary Oberije heeft een aantal jaren geleden een classificatie taak in bestraalde longkankerpatiënten gedaan die voor mij de reden en in zekere zin ook de start vormt van het onderzoek waarover ik u vandaag wat zal vertellen. De test die Cary uitvoerde was dat zij MAASTRO Clinic dokters een stapel dossiers gaf van longkankerpatiënten die in het verleden door ons waren bestraald. Daarin stond alles wat bekend was op het moment dat deze patiënten begonnen met de bestraling, wat er daarna gebeurde had ze verwijderd. En ze vroeg aan deze dokters welke van de patiënten na twee jaar nog in leven waren. Ze moesten de mensen waarvan ze zeker waren dat ze nog in leven zouden zijn aan de linkerkant van de tafel leggen, en de mensen waarvan ze zeker waren dat ze binnen twee jaar aan hun kanker zouden overlijden aan de rechterkant, en de mensen waarover ze twijfelden in het midden van de tafel.

U zult begrijpen dat dit eenzelfde soort experiment is als dat van die tulpen en we kunnen daarom opnieuw de AUC bepalen van deze voorspelling door onze dokters. Wie denkt dat de MAASTRO dokters daar heel goed in zijn, dat zij een AUC hebben die dicht bij 1 ligt? Wie denk dat de MAASTRO dokters dat helemaal niet kunnen, dat ze een AUC hebben die dicht bij 0.5 ligt? En wie denkt dat ze dat een beetje kunnen, dat ze een AUC in de buurt van 0.7 hebben?

Als ik dit vraag antwoorden de meeste mensen 0.7 of hoger. Maar de realiteit is dat onze dokters in die test 0.57 scoorden. En als je daar met statistiek naar kijkt, kon je dat niet onderscheiden van 0.5. De voorspellingen werden gedaan door dokters die net uit hun opleiding kwamen en door dokters die jarenlange ervaring hadden. Dat maakte niets uit.

Je kan wel kritiek hebben op deze studie, het ging immers op patiënten die dokters niet zelf gezien hadden, ze moesten het doen met het dossier van historische patiënten. Wellicht zouden de voorspellingen beter zijn in de patiënten van vandaag. Daarom hebben we die studie herhaald als een prospectieve studie, waarbij we de dokters vroegen om de voorspelling te doen in patiënten die ze zelf daadwerkelijk zagen. En ze moesten niet alleen overleving maar ook bijwerkingen van radiotherapie voor longkanker voorspellen. Typische bijwerkingen zijn kortademigheid omdat je de gezonde longen voor een stuk mee bestraald en moeite met slikken en eten omdat je de slokdarm ook niet altijd kan ontwijken. Het resultaat van deze prospectieve studie? Een AUC van 0.56 voor overleving en niet veel beter voor andere uitkomsten zoals bijwerkingen.

Hebben we bij MAASTRO zo'n slechte longkanker dokters? Nee, meerdere studies hebben dit nu aangetoond ook voor andere kankers in andere instituten. Deze zomer bijvoorbeeld op het grootste congres voor radiotherapie werd eenzelfde studie gepresenteerd in palliatieve radiotherapie in Stanford California. Palliatieve radiotherapie geven we in patiënten die we niet meer kunnen genezen. Vaak om klachten zoals pijn te verminderen. Dokters uit Stanford bleken in deze patiëntengroep overleving ook heel slecht te kunnen voorspellen, ze waren vaak veel te optimistisch. Als je als patiënt bij MAASTRO of bij Stanford of waar dan ook binnenloopt en vraagt "Dokter leef ik nog over een jaar?" dan kun je net zo goed een muntje opgooien.

Hoe komt dat nou, dat dokters niet goed kunnen voorspellen wat de uitkomsten zullen zijn van hun behandeling? Eigenlijk weten we dat niet goed, maar we denken dat de volgende zaken een rol spelen.

Ten eerste verzuipt een dokter momenteel in gegevens over patiënten. Beeldvorming zoals CT en MRI scans, waardes bepaald uit het bloed, genetische eigenschappen van de tumor, hoe een patiënt zich voelt, de stappenteller aan hun pols. Het zijn allemaal gegevens die misschien de uitkomst van een behandeling bepalen. Maar uit studies over het menselijk denkvermogen weten we dat mensen, en dokter dus ook, maar 5 gegevens verwerken als ze een besluit moeten nemen. Meer dan 5 gegevens gaat je cognitieve capaciteiten als mens te boven. Maar dokters zien er tientallen, honderden gegevens en dat worden er steeds meer. Het is dus niet zo raar dat ze dat niet kunnen verwerken als ze een besluit moeten adviseren aan een patiënt. Er is een gegevens explosie gaande ook wel de data tsunami genoemd.

Onze medisch handelen, hoe we mensen opleiden en hoe we innoveren, is sterk gebaseerd op een concept genaamd "evidence based medicine", geneeskunde op basis van bewijs. De hypothese van evidence based medicine is dat we, op het kruispunt van de gegevens van een patiënt en de mogelijke besluiten, gebruik kunnen maken van het best beschikbare bewijs van in de vorm van gepubliceerde, goede wetenschappelijk studies en richtlijnen. Maar de realiteit is dat niemand de wetenschappelijke literatuur kan bijhouden. Als je dat bijvoorbeeld voor longkanker en radiotherapie wil doen, moet je als dokter zo'n 4 uur per dag artikelen lezen, 365 dagen van het jaar. Dat kan dus niet en de hoeveelheid literatuur is ook nog eens sterk aan het stijgen...

Daarnaast is de literatuur sterk vooringenomen, het bewijs dat je leest komt van maar 3% van de patiënten, want dat is het aantal kankerpatiënten dat in studies zit. Deze 3% is sterk geselecteerd. Ouderen, mensen met, naast kanker, andere ziektes worden zelden in studies geïncludeerd. Maar die zie je als dokter wel elke dag in het ziekenhuis. En dus kun je je afvragen of de resultaten van die studies wel van toepassing is op de patiënt die voor je zit. Vaak zijn goede studies niet goed mogelijk,

zeker als het technologie betreft. Als je bijvoorbeeld, zoals Jacques Borger en Liesbeth Boersma, een studie doet naar het effect radiotherapie op het hart in borstkanker patiënten moet je 10 jaar wachten totdat je de resultaten van de studie weet, omdat hartproblemen zich vaak pas na 10 jaar openbaren. Maar de afgelopen 10 jaar is radiotherapie drastisch veranderd, dus wat moeten we met het bewijs uit een dergelijke studie voor de huidige tijd? Medicijnen veranderen niet, paracetamol blijft paracetamol, maar technologie verandert continu.

Een vooraanstaand tijdschrift, de British Medical Journal, heeft een paar jaar geleden 3000 veelvoorkomende behandelingen geselecteerd om te kijken welk bewijs er was daarvoor. Voor 4 van de 10 behandelingen is er enig bewijs dat ze werken. Voor 1 van de 10 is er bewijs dat ze niet werken maar we doen ze toch. Voor de helft van de behandelingen is er helemaal geen bewijs. Niet dat ze werken en ook niet dat ze niet werken.

En dus vragen we de dokter terwijl ze verzuipen in gegevens, in behandelopties en in bewijs dat niet relevant is, te voorspellen wat de uitkomst van een behandeling zal zijn. Vindt u het raar dat blijkt dat ze dat niet goed kunnen? Maar als ze niet de uitkomst kunnen voorspellen van een behandeling in een individuele patient, hoe kunnen ze dan adviseren of behandeling A beter is dan B?

Het lange termijn doel van mijn onderzoek en de leerstoel "Clinical Data Science" die ik vandaag aanvaard met deze oratie, is om van geneeskunde op basis van bewijs meer naar een geneeskunde op basis van data te komen, van "evidence based medicine" naar "data driven medicine". Waarin we de dokter en patiënt helpen om besluiten te nemen op basis van gevalideerde voorspellingen. Voorspellingen die relevant zijn voor de data van de patiënt die voor de dokter zit. En die geleerd zijn van de data van andere mensen die hun voor zijn gegaan. Leren van andermans data.

Hoe zie ik dat dan voor me, dat "data driven medicine". Nou stel dat je gisteren alle data van alle kankerpatiënten op de hele wereld die ooit behandeld zijn, hebt verzameld. De nacht heb je gebruikt om te leren van deze data, daar kennis uithalen. En die kennis gaan we vandaag gebruiken om samen met een patiënt een besluit te nemen. Vervolgens noteer je wat de patiënt gekozen heeft en wat het resultaat van die keuze was en dan heb je weer nieuwe data waarvan je vannacht weer kan leren. Kunnen we naar een situatie dat we elke 24 uur iets leren van alle patiënten die we hebben behandeld in plaats van te moeten leunen op bewijs van 3% van de patiënten die 15 jaar geleden zijn behandeld? In plaats het onmenselijke te vragen van onze dokters als ze advies aan patiënten moeten geven. Dat is de vraag die ik me stel in deze leerstoel "Clinical Data Science".

Leren van andermans data – hoe zorg je dat je kan leren van andermans data

Laten we beginnen met gisteren, hoe zorg je dat je alle data van alle patiënten die ooit behandeld zijn in de wereld beschikbaar krijgt om van te leren? Hoeveel is dat eigenlijk? Laten we even een sigarendoos berekening doen. Jaarlijks krijgen zo'n 14 miljoen mensen kanker in de wereld. Stel dat we data van de afgelopen 10 jaar zouden willen hebben dan heb je het dus over 140 miljoen patiënten. De hoeveelheid data varieert natuurlijk, in westerse ziekenhuizen genereren we al snel 10 gigabyte aan data per patiënt, in minder ontwikkelde landen minder, stel 0.1 gigabyte. Dan ligt het totale datavolume wat we beschikbaar moeten krijgen ergens tussen de 14 en de 1400 petabyte, laten we zeggen 500 petabyte. Is dat veel? Ja, maar niet onoverkomelijk veel. Alhoewel het een bedrijfsgeheim is, wordt uit het energieverbruik van Google geschat dat zij 30X zoveel data hebben. Bij CERN in Geneve hadden ze een jaar geleden 530 petabyte aan data staan, geproduceerd door hun deeltjesversneller. De hoeveelheid data is dus niet het grootste probleem. Wat een groter probleem is, is dat het overgrote deel (80%) van de data niet gestructureerd zijn. Het is bijvoorbeeld gedicteerde tekst of beelden waar je eerst allerlei technologie op moet loslaten om daar structuur in te vinden zodat je er wat van kan leren. Vertalingen, "natural language processing", IBM Watson en technieken zoals Siri op je iPhone die uit een tekst snappen wat er moet gebeuren zijn voorbeelden van opkomende technologie die steeds beter wordt in structureren en vervolgens snappen van tekst.

Voor beelden is dat ook sterk in opkomst. Misschien heeft u wel eens een Google auto gezien met zo'n camera op het dak die fotos maakt van de straat voor gebruik in google Maps. Streetview heet dat. Met een techniek genaamd deep learning kan je nu een willekeurig foto aan Google geven en Google kan dan vertellen waar op de wereld die foto genomen is. Apple haar gezichtsherkenning is een ander voorbeeld van de mogelijkheden tot structureren van beelden.

In kanker hebben we heel veel afbeeldingen van patiënten en hun tumoren zoals PET, CT en MRI beelden. Bij MAASTRO hebben we een lange traditie om deze beelden te analyseren en te structureren. Ik heb het genoeg gehad diverse promovendi als co-promotor te hebben begeleid die allemaal een bijdrage hebben geleverd aan een veld dat we nu Radiomics noemen. Met Radiomics berekenen we automatisch, uit beelden die we elke dag maken van patiënten, honderden kenmerken van tumoren. Denk dan aan de grootte van een tumor, of de vorm ervan. We maken van de beelden dus gestructureerde informatie.

Als de hoeveelheid data niet echt een probleem is en er technologieën opkomen om het gebrek aan structuur op te lossen, wat is dan wel het probleem?

Het belangrijkste probleem is dat die data van die 140 miljoen patiënten verspreid staat over 100.000 ziekenhuizen. Als je een ziekenhuis vraagt om een bepaalde dataset te delen, moet je over het algemeen een of meer van de volgende problemen overwinnen.

1. Ik heb geen tijd om de data van mijn patiënten te delen. Ik heb al geen tijd om al die data te noteren. En als ik de data wel heb genoteerd in mijn dossier, heb ik geen tijd om door al mijn dossiers te gaan en de dataset netjes samen te stellen. We noemen dit het administratieve probleem.
2. Ik wil de data van mijn patiënten niet delen. Want als jij mijn data hebt, ontdek je misschien dat ik iets niet goed doe in mijn patiënten en als dat naar buiten komt verlies ik patiënten

aan het ziekenhuis even verderop, krijg ik een slechte naam etc. Of als jij mijn data hebt en daarover een publicatie schrijft, dan kunnen mijn promovendi niet meer over diezelfde data publiceren en dan kunnen ze niet promoveren. We noemen dit het politieke probleem.

3. Ik mag de data van mijn patiënten niet delen. De wetgeving in mijn land verbiedt het mij om data te delen. Of ik moet de privacy van mijn patiënten beschermen. Het ethisch-juridische probleem. Wordt overigens vaak als een drogreden gebruikt voor geen tijd en niet willen. In elk land dat ik ken, is het mogelijk om medische data met elkaar te delen om daar onderzoek mee toe doen mits je het op een goede manier doet, bijvoorbeeld door de data te anonimiseren. David Townsend heeft recent op deze Universiteit zijn inaugurele rede gehouden op dit onderwerp en ik kijk uit met hem samen te werken om deze barrière te slechten.

Met zoveel problemen om data met elkaar te delen vraagt u zich wellicht af hoe men het toch voor elkaar krijgt om voor onderzoek data met elkaar te delen. De eerste methode is het opzetten van een klinische studie. Vaak gesubsidieerd door de Europese Unie, de Nederlandse overheid of organisaties zoals KWF Kankerbestrijding, heeft een onderzoeker voor een specifieke onderzoeksvraag tijd en geld om toestemming te kunnen vragen aan patiënten, om lokale datamanagers te betalen om alle gewenste data netjes te noteren en om deze in een centrale database op te slaan. Dat kost heel veel geld en daarom zitten slechts 3 % van de patiënten in klinische studies, zoals eerder gezegd.

De tweede methode is het opzetten van een centrale registratie, zoals de Nederlandse kankerregistratie, waarin weliswaar gegevens van alle Nederlandse kankerpatiënten staan maar waar maar per patiënt maar een zeer beperkt stuk van de gegevens in staan omdat het anders niet te doen is om dat te verzamelen. Er staat bijvoorbeeld nauwelijks iets in over radiotherapie, geen beelden, geen genetische eigenschappen etc.

De derde en nieuwste methode is je data toevertrouwen aan een grote speler zoals Apple's HealthKit, IBM's HealthCloud of Philips' HealthSuite. Zij maken over het algemeen een kopie van al jouw data, centraliseren die, tegenwoordig vaak in de cloud. Ze proberen dan wat uit de data te leren en die kennis vervolgens te gelde te maken. Dat stuit, vooral in Europa, op nogal wat bezwaren omdat we niet goed weten wat die grote bedrijven allemaal met die data doen en of we de kennis die ze uit onze data halen wel voor een redelijk bedrag terugkrijgen. Krijgen we commerciële silo's van kennis op gezondheidszorg?

Ik kom uit een ingenieursgeslacht. Ik ben ingenieur, mijn twee zussen zijn ingenieurs, mijn vader en moeder zijn ingenieur, mijn opa had een garage, mijn andere opa werkte bij de sodafabriek in Herten en mijn oudoom, hier aanwezig, was treinmachinist in de DDR. En als ingenieur zijn we de afgelopen jaren bezig geweest een radicaal andere oplossing te maken.

Zoals vaker begon dat bij het stellen van de juiste vraag. In ons geval: Kunnen we van elkaars data leren zonder dat we de data met elkaar delen? Als de data het ziekenhuis niet verlaat heb je geen privacy probleem, hoef je data niet meer uit je dossiers te halen, hou je als ziekenhuis controle over wat er met de patiënten data gebeurt en heb je veel minder technische issues omdat je grote hoeveelheden data niet meer hoeft te verplaatsen.

Kunnen we van andermans data leren zonder dat we de data verplaatsen? Als de berg niet tot Mohammed komt, zal Mohammed tot de berg gaan (in het Duits “Wenn der Berg nicht zum Propheten kommt, muss der Prophet zum Berg kommen.”). Als we de data niet naar het leren kunnen verplaatsen moet het leren naar de data.

Onze oplossing daarvoor is om de vraag die we aan de data willen stellen te verpakken in een computerapplicatie, een stukje software, en deze naar het ziekenhuis te sturen. Deze applicatie stelt de vraag aan de lokale database, bijvoorbeeld “welke factoren bepalen in de patiënten van dit ziekenhuis de overleving van longkankerpatiënten na een bestraling”. De applicatie deelt vervolgens het antwoord met de buitenwereld maar deelt niet de data.

Dat principe van het leren naar de data brengen klinkt eenvoudig maar we zijn daar de afgelopen jaren flink aan het puzzelen geweest om dat voor elkaar te krijgen met een grote groep promovendi, postdocs en senior onderzoekers teveel mensen om te noemen, maar veel van hun zitten in de zaal.

Aan de ene kant is er een technologische uitdaging. Hoe distribueer je applicaties over heel de wereld, hoe zorg je dat die veilig zijn en geen virussen bevatten, hoe zorg je dat ze open zijn, dat je als ziekenhuis weet welke data gebruiken en wat met de buitenwereld communiceren.

Er is ook een wiskundige uitdaging vooral omdat je de kennis in brokjes binnenkrijgt. Hoe combineer je kennis uit diverse ziekenhuizen op een goede manier. Is de kennis die je met distributed learning opdoet niet minder goed dan gewoon alle data op een plek samenbrengen?

Tenslotte is er een ontologische uitdaging, hoe zorg je dat data over heel de wereld dezelfde betekenis heeft. Als je de data niet meer met elkaar deelt, heb je niet meer de luxe van menselijke interpretatie van de data. Als voorbeeld, die applicatie spreekt geen Chinees of Nederlands dus moet je de data taalafhankelijk maken anders kan een computer daar niet van leren.

We hebben veel voortgang geboekt de afgelopen jaren in deze technologische, wiskundige en ontologische uitdagingen. Inmiddels hebben we een lerend netwerk, waarbij 20 centra uit bijna heel de wereld, op de door ons ontwikkelde manier, van elkaar data kunnen leren. Australië, Azië, Afrika, Europa, Noord Amerika. Allemaal doen ze mee of hebben ze gezegd mee te gaan doen op korte termijn.

Over dit andermans data gedeelte kan ik nog uren doorpraten, omdat er zoveel interessante aspecten aan zitten. Maar ik zal stoppen. Maar niet voordat ik u heb geschetst waar we de komende jaren in gaan investeren. Enerzijds is dat het groeien van ons wereldwijde netwerk, vooral door het makkelijker te maken voor nieuwe centra om zich aan te sluiten, we moeten immers van 20 naar 100.000 ziekenhuizen groeien.

Maar we richten ons ook steeds meer op Nederland, ons concept van research naar de data brengen is een van de elementen in een nationaal initiatief genaamd Health-RI, Health Research Infrastructure. Ondersteund door alle universitaire ziekenhuizen, de Koninklijke Nederlands Akademie van Wetenschappen en vele andere belanghebbenden is dit een gezamenlijk voorstel om te komen tot één nationale onderzoeksinfrastructuur voor medische data waarbij Nederland als gidsland dient voor de EU.

Limburg heeft een voortrekkersrol. Ondersteund door de provincie en in een samenwerking tussen de Universiteit Maastricht en Hogeschool Zuyd, ontwikkelen we in Limburg een voorbeeld van een dergelijke infrastructuur. Dit heet de Personal Health Train en hierin krijgen zowel Limburgers meer controle over wat er met hun eigen data gebeurt maar kunnen tegelijkertijd onderzoekers ook makkelijker leren van de beschikbare data. In ons regionale oncologie netwerk Oncozon worden ook vele initiatieven ontplooid om te zorgen dat we op een goede manier van elkaars data kunnen leren om de kankerzorg regionaal beter te krijgen.

In Maastricht investeert ook de Universiteit en het MUMC+ al flink in data science en een research data infrastructuur voor patiënten data. Michel Dumontier is recent benoemd als universiteitshoogleraar en met zijn komst, de oprichting van een instituut voor data science aan deze universiteit en het research IT programma van het ziekenhuis hebben we prachtige kansen om hier de komende jaren echt meters te maken.

Opvallend is dat we in ons onderzoek op een globale schaal begonnen zijn en dat onze ideeën nu pas op nationaal, regionaal en lokaal acceptatie krijgen. Leren van andermans data is soms makkelijker dan leren van je eigen data.

Leren van andermans data – hoe leer je van data

Tot zover data. Ik heb tot nu toe nogal makkelijk gedaan over het leren. Je maakt een lerende applicatie en je stuurt die rond. Maar wat wil je eigenlijk leren, wat gebeurt er in een lerende applicatie en hoe zorg je dat wat je leert klopt?

Wat we willen leren zijn zogenaamde predictieve modellen. Modellen die voorspellen wat de uitkomst zal zijn van een bepaalde behandeling in een bepaalde patiënt. Even terug naar de tulpen en de voorspellingen van dokters, we willen uit al die data beter kunnen voorspellen dan dokters. Ons doel is een AUC van ongeveer 0.85. Wat we willen voorspellen ligt aan de ziekte, voor longkanker is dat overleving, kortademigheid en slikklachten. Recent was ik in Toronto en daar zijn meer geïnteresseerd in hartfalen. Dat bracht me terug naar mijn promotie bij cardiothoracale chirurgie die ook voor een belangrijk deel over hartfalen ging.

Een voorspellend model kun je leren uit data. Dat is het terrein van machine learning of data mining. Ik ga niet de voor- en nadelen van diverse machine learning technieken aan u uitleggen. Random forest, neural network, deep learning, er zijn er nog 100 en één.

Interessantere vragen zijn of we genoeg data hebben om de computer een goed voorspellend model te laten leren. Kunnen we de data gewoon aan een computer geven en die genereert dan wel de kennis voor ons. Dat een computer geprogrammeerd kan worden om dat te kunnen is een stiekeme droom voor veel nerds zoals ingenieurs en misschien wel een nachtmerrie voor vele anderen.

Om te weten of dat realistisch is, moeten we even langs Itteren en Borgharen. Voor de mensen die niet van Maastricht zijn, dat zijn twee kerkdorpen even ten noorden van Maastricht. Die dorpen overstroomden regelmatig, gemiddeld eens per 50 jaar, door de Maaswerken die nu gaande zijn moet dat worden teruggebracht naar eens per 250 jaar. Het idee dat Nederland bang moet zijn voor overstromingen van de zee omdat er een groot deel in het westen onder zeeniveau ligt, is niet echt terecht. De kans op overstromingen in Zeeland is bijvoorbeeld maar eens per 4.000 jaar en in Holland eens per 10.000 jaar. Daar zouden Borgharen en Itteren zo voor tekenen. Maar hoe bepaal je dat eigenlijk, eens per 10.000 jaar? 10.000 jaar geleden liepen er in Itteren rendieren rond, stonden er geen tulpen in Holland, en mensen konden niet eens schrijven, laat staan dat men waterstanden noteerde. De theorie om dat toch te kunnen inschatten, heet de "Extreme Value Theory". Daaruit bepaalt men uit de variatie een beperkt aantal waarnemingen, bijvoorbeeld de waterstanden van de laatste 100 jaar, wat een extreme waarde in waterstanden zou kunnen zijn als je 10.000 jaar vooruitkijkt.

Een parallel is ons werk. Stel je laat een computer nerd van 100 patiënten een aantal malen een voorspellend model leren. Uit de variatie in AUC oftewel voorspellend vermogen kun je dan bepalen of de computer een perfect model, een AUC van 1.00, kan leren als die 1 miljoen patiënten ter beschikking heeft om van te leren. Uit onze analyses tot nu toe lijkt een computer dat niet te gaan lukken. Er zijn waarschijnlijk niet genoeg data in de wereld van voldoende kwaliteit over kankerpatiënten beschikbaar om automatisch een perfect model van te leren. De technologieën die machine learners op dit moment gebruiken bijvoorbeeld in het eerder genoemde voorbeeld waarbij Google een locatie bij een willekeurige foto kan bepalen maakte gebruik van 126 miljoen plaatjes. Zoveel longkankerpatiënten zijn er niet en de variatie tussen patiënten is waarschijnlijk te groot.

De oplossing zit hem wat ons betreft in het koppelen van de machine learning nerd met zijn of haar computer aan iemand die de ziekte, data en de relaties daartussen begrijpt. Dat noem je in machine learning dialect een domain expert maar in ons geval heet zo'n persoon gewoon een dokter. De combinatie van expertise creëert niet alleen betere voorspellende modellen maar ook modellen die beter geaccepteerd worden door dokter omdat ze "logisch" zijn en geen "black box". De dokter kan helemaal niet vervangen worden door een computer, maar wel geholpen worden. Het samen met dokters leren van logische modellen voor kanker en misschien wel hartfalen is een belangrijk doel voor de komende jaren.

Als je eenmaal een voorspellend model hebt geleerd dat een beetje logisch eruit ziet, is de volgende vraag hoe goed is het model? De validatie van een model mag je nooit doen in de data waaruit het model is geleerd. Dit mag niet omdat het model precies passend is gemaakt op de data waarvan het is geleerd. Beter is het om het model te valideren in een dataset die van buiten komt, bijvoorbeeld van een ander ziekenhuis. In de data infrastructuur die ik eerder heb beschreven is dat vrij eenvoudig, je stopt het model in een de applicatie, je stuurt die naar een ander ziekenhuis, je vergelijkt de voorspellingen van het model met de werkelijke uitkomsten. Daar kun je een AUC uit bepalen net als met die tulpen. Dat heet externe validatie en zou wat ons betreft een voorwaarde moeten zijn om modellen te mogen publiceren. Veel modellen falen bij externe validatie, ze zijn veel te veel passend gemaakt op de eigen data, een probleem genaamd "overfitting".

Om u een voorbeeld te geven, we hebben een, vrij eenvoudig, longkanker model geleerd in MAASTRO Clinic data. Dat model kan overleving voorspellen in longkankerpatiënten bestraald met een hoge dosis radiotherapie, een curatieve dosis. Vervolgens hebben we dat model naar Australia gebracht en daar gevalideerd in de data in Sydney. De AUC van dat model was 0.69. Niet super maar wel al veel beter dan de 0.57 van dokters. De race is om de AUC omhoog te krijgen en modellen te leren voor meer kankers, andere ziektes zoals hartfalen en meer uitkomsten. Die toekomst ziet er behoorlijk rooskleurig uit in dit veld. Er komt steeds meer en betere data beschikbaar en de machine learning algoritmes worden beter. Ik heb er alle vertrouwen in dat we voor de meeste uitkomsten op een AUC van 0.85 of hoger gaan komen met vereende krachten van de data mensen, de machine learners en de dokters.

Leren van andermans data – welke effecten heeft dat

In het laatste deel van mijn oratie, wil ik met u praten over de maatschappelijke relevantie. Misschien heeft u tot nu toe het gevoel dat u luistert naar een abstract verhaal over data en leren. Maar wat heeft het voor zin? Worden daar mensen beter van? We moeten toch met zijn allen kanker genezen, is de investering in onderzoek naar nieuwe medicijnen niet veel relevanter?

Ik heb het grootste gedeelte van mijn jeugd in Zwolle gewoon. En tegenover mij woonde mijn vriendje, die trouwens ook ingenieur is geworden. Ik ging vaker bij hem op zondag studio sport kijken, samen met zijn drie broers en hun moeder, zij waren voor Ajax en ik voor PSV. Hun moeder keek altijd mee en soms kregen we een bakje chips van haar. 20 jaar later zag ik haar weer op het bordes van Chateau Neercanne. Op de receptie van haar oudste zoon die zojuist was getrouwd met een Maastrichste schone. Ze zag er niet goed uit, ze had volgens mij een pruik op, was verzwakt, haar man had haar arm vast. Ze was misselijk, ik denk niet dat ze chips had kunnen binnenhouden. Niet lang daarna hoorde ik dat ze overleden was, aan longkanker.

Ik moest aan de moeder van mijn vriendje denken nadat ik in Australië was geweest. In Sydney. In tegenstelling tot in Maastricht zien zij in Sydney longkanker lang niet altijd als een mogelijk te genezen ziekte. In Maastricht (en in heel Nederland) hebben we een richtlijn dat we alle patiënten een hoge dosis geven, we gaan er in Nederland vanuit dat we patiënten nog kunnen genezen of in ieder geval hun leven kunnen verlengen, we geven een zogenaamde curatieve of radicale behandeling. In de helft van de patiënten besluit de Australische dokters echter tot een palliatieve behandeling, ze geven dan een lage dosis bestraling om klachten te bestrijden maar zijn niet uit op genezing of levensverlenging. De overleving van die palliatief behandelde patiënten is 20% na twee jaar, slechts één op de vijf is nog in leven na twee jaar. Maar dat besluit van die Australische dokters is subjectief en we weten uit het begin van deze oratie dat dokters eigenlijk niet weten wat de uitkomst in patiënten zal zijn. Zo zijn er sommige patiënten die de Australische dokter selecteert die misschien beter een hoge dosis had kunnen geven. Ons model dat we gevalideerd hebben in Sydney kan dat voorspellen en het laat zien dat er patiënten zijn wiens overlevingskansen zouden stijgen van 20% naar 60% na twee jaar. Een drie keer zo hoge kans op leven als ze die patiënten een hoge dosis zouden hebben gegeven zoals we doen in Nederland. Nu denkt u misschien dat het wel fijn is om in Nederland te zijn en niet in Australië en voor sommige patiënten is dat waar. Maar er zijn ook patiënten die wij in Nederland met een hoge dosis behandelen en waarvan ons model zegt dat dat nutteloos is, de overleving daarvan is niet anders dan in Australië waar in deze patiënten een veel minder intensieve behandeling geven. In Nederland zijn er patiënten die nog maar 6 maanden te leven hebben en waarvan wij ze 2 maanden doodziek maken met onze intensieve behandelingen en deze patiënten zijn waarschijnlijk in Australië beter af.

Bij welke groep hoorde de moeder van mijn vriendje? Als haar de keuze was gegeven tussen een agressieve behandeling met een wat langere overleving maar wel veel bijwerkingen versus wat minder lang leven maar wel met haar eigen haar en niet misselijk op het bordes in Neercanne chips kunnen eten. Wat zou ze hebben gekozen?

De verwachting is dus dat het gebruik maken van modellen in sommige patiënten dus leidt tot een veel grotere kans op overleving en in sommige patiënten tot het bijna geheel voorkomen van bijwerkingen, eenvoudig weg door beter te kiezen in een individuele patiënt. Er is geen medicijn of technologie in ontwikkeling dat dergelijke grote verwachte effecten op overleving en bijwerkingen

heeft. Niet immunotherapie, niet protonen therapie, en niet alle andere therapieën waarover u misschien heeft gehoord of gelezen.

Ik dacht tot niet zo lang geleden met dat dergelijke verwachtingen de introductie van voorspellende modellen een eitje zou zijn. Ik ben daar inmiddels van teruggekomen. De introductie van voorspellende modellen, ook wel “decision support systems” in de kliniek is een totaal ander uitdaging dan ik initieel verwachtte. Modellen gebruiken kost tijd van de dokter en die is er niet. Sommige modellen zijn ook best complex, een black-box. Er zit ook een communicatiekloof tussen onderzoekers (computer nerds & ingenieurs) en de dokters die het moeten gebruiken. Die modellen geven ook de indruk niets nieuws op te leveren. Wat je doet met modellen is beter selecteren tussen bestaande behandelingen. Voor veel mensen (zowel patiënten als dokters) en ook subsidiegevers is dat niet sexy. Je belooft immers niet kanker te genezen. Er is kritiek op de kwaliteit van de modellen. Hoe zijn die geleerd, kloppen die wel, zijn de laatste inzichten erin verwerkt.

Twee dagen geleden hadden we een bezoeker uit Denemarken die daar wat interessants over zijn. Hij noemde het een asymmetrie in verwachtingen. Hij gaf het voorbeeld, stel dat je in de krant leest “man overleden na auto-ongeluk in de Verenigde Staten”. Je zou zeggen waarom is dat nieuws, mensen maken fouten en soms met tragische afloop? Maar als je doorleest staat er “man overleden na auto-ongeluk in de Verenigde Staten terwijl hij in een Tesla reed met auto-pilot”. Dat vinden we erg. We verwachten dat technologie perfect is terwijl we dat niet van mensen verwachten. Modellen zijn aantoonbaar beter dan dokters maar niet perfect en daarom niet goed genoeg. Asymmetrie van verwachtingen over kwaliteit.

Als het moeilijk is om het bij de dokters geaccepteerd te krijgen, moeten we ons misschien richten op patiënten. Het zogenaamde patient decision aids, shared decision making, samen beslissen, de patiënt moet aan zet, regie hebben. Dat lijkt aantrekkelijk, de modellen die wij leren inzetten om mensen zelf te laten beslissen, bijv. als app op je telefoon of met een website. En bij MAASTRO zijn we daar ok zeker mee bezig. We moeten af van de paternalistische dokter die het wel voor de patiënt besluit. Een zorgverzekeraar als CZ ziet dit ook als de toekomst.

Maar kan een kankerpatiënt wel de informatie tot zich nemen en een afgewogen besluit nemen? Helemaal in een ziekte zoals in longkanker waar tussen het moment dat je hoort dat je longkanker hebt en je moet starten met de chemo vaak slechts dagen ligt. Dit omdat je zo snel mogelijk moet behandelen als je de kans op genezing zo hoog mogelijk wil hebben. Longartsen waarschuwen ervoor dat samen beslissen niet gaat werken in de groep patiënten die zij zien. In mei was hier in deze zaal een oratie van prof. Rademakers over dit onderwerp, waarin zij de actieve patiënt als utopie bestempeld.

De bijdrage van mijn leerstoel aan deze complexe materie is noodgedwongen beperkt. Ten eerste wil ik ervoor zorgen dat we de beste mogelijke data van en voor iedereen ter beschikking hebben zodat wat we ook doen, dat in iedere geval gestoeld is op harde gegevens en geëvalueerd kan worden. Daarnaast zullen we doorgaan met continue betere modellen leren, zodat er in ieder geval aan de kwaliteit van modellen steeds minder wordt getwijfeld. Tenslotte zullen we opportunistisch gebruik maken van de kansen die ons geboden worden. Zo’n kans is protonentherapie, dat we ook hier in Maastricht in ZON-PTC gaan krijgen. Daarin hebben we met zijn allen in Nederland besloten om van elkaars data modellen te leren en deze modellen leidend te laten om te bepalen wie er in aanmerking komt voor protonen. De introductie van protonen therapie in Nederland is een

uitgelezen kans om te laten zien wat je met leren van andermans data kan bereiken. Een prachtige uitdaging gedurende de komende jaren voor deze leerstoel.

Dankwoord

Tot slot het dankwoord. Dat zal ik conform het vademecum algemeen en kort houden. Dit hoogleraarschap is een waardering voor het werk dat vele mensen bij MAASTRO Clinic, de universiteit en het ziekenhuis de afgelopen jaren hebben gedaan. Ik heb gedurende de oratie al diverse groepen en mensen genoemd en ik voel me vaak niet meer dan de cheerleader voor hen die het echte werk doen. Leren van andermans data is voor mij ook strijken met andermans veren. Ik dank hun daarvoor en kijk uit naar de komende jaren waarin we door zullen gaan met dit onderzoek.

Een persoon heb ik nog niet genoemd. Philippe Lambin is de drijvende kracht achter het onderzoek bij MAASTRO en zonder twijfel de beste mentor voor mijzelf die ik me had kunnen wensen de afgelopen 13 jaar. Het onderzoek dat ik doe en dit hoogleraarschap is in hoge mate aan hem te danken. Philippe, ik en de andere vier hoogleraren van MAASTRO zullen van elkaar blijven leren en onze krachten bundelen om top kankeronderzoek te doen.

Ik sta hier ook voor vrienden en familie. Mijn schoonouders op wie ik altijd kan rekenen. Mijn ouders die mij altijd hebben gesteund. Zonder hen had ik dit nooit bereikt. Tenslotte zitten mijn vrouw en kinderen ook in de zaal. Ik houd werk en privé nogal gescheiden. Zij weten eigenlijk niet wat ik doe (misschien vandaag iets meer) en dat hoeft ook niet. Zij willen alleen weten wanneer ik er ben en eisen dan mijn aandacht op. En dat is prima. Hun danken voor hun bijdrage aan wat ik vandaag heb verteld, zou nep zijn, zij hebben daar weinig aan bijgedragen. Maar recent zij iemand tegen mij die veel met hoogleraren omgaat: "In mijn ervaring is slechts 10% van de hoogleraren nog redelijk normaal. Jij lijkt daar een van te zijn." Daar dank ik hun wel voor.

Dank voor uw aandacht. Ik heb gezegd.