

NO FAIRNESS WITHOUT AWARENESS

Toegepast onderzoek naar kansengelijkheid in het hoger onderwijs
Intreerede lectoraat Learning Technology & Analytics

DR. T.C. BAKKER



NO FAIRNESS WITHOUT AWARENESS

Toegepast onderzoek naar kansengelijkheid in het hoger onderwijs
Intreerede lectoraat Learning Technology & Analytics

DR. T.C. BAKKER

Het onderzoek uit deze intreerede is uitgevoerd door het lectoraat Learning Technology & Analytics van het Centre of Expertise Global & Inclusive Learning en de Faculteit IT & Design van De Haagse Hogeschool.

DOI: <https://doi.org/10.5281/zenodo.14204674>

Delen van deze intreerede zijn uitgesproken op 21 november 2024 aan De Haagse Hogeschool. De bijeenkomst is in zijn geheel online na te zien op de website van De Haagse Hogeschool.

Omslag: *Een zondagmiddag op het eiland La Grande Jatte*, George Seurat, bron: Wikipedia

© Dr. T.C. Bakker, Den Haag, 2024

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij mechanisch, door fotokopieën, opnamen, of enige andere manier, zonder voorafgaande schriftelijke toestemming van de auteur.

DE HAAGSE HOGESCHOOL

NO FAIRNESS WITHOUT AWARENESS

Toegepast onderzoek naar kansengelijkheid in het hoger onderwijs
Intreerede lectoraat Learning Technology & Analytics

door

Dr. T.C. Bakker

*Footfalls echo in the memory
Down the passage we did not take
Towards the door we never opened.*

— T.S. Elliot

The accommodation of differences is the essence of true equality

— Will Kymlicka

Opgedragen aan
Coby Breur, 1941–2013,
en Gerrit Bakker, 1936–2017

INHOUD

I	Inleiding – Onderzoek naar kansengelijkheid en fairness – Conclusies	
1	Inleiding	1
1.1	Learning analytics, Educational Data Mining en sociale rechtvaardigheid	2
1.2	Doel en opbouw introereede	5
2	Kansengelijkheid in het hoger onderwijs	7
2.1	Persoonlijke motivatie voor kansengelijkheid	7
2.2	Het belang van hoger onderwijs voor kansengelijkheid	8
2.3	Beperkingen in het denken over kansengelijkheid en hoger onderwijs .	9
3	Analyse van bias en fairness in studiedata	11
3.1	Diversiteit in instroom en uitstroom	11
3.2	Zorgen over kansengelijkheid en bias in studiedata	12
3.3	Sensitieve kenmerken versus bijzondere persoonsgegevens	14
3.4	Huidig onderzoek naar fairness in het (hoger) onderwijs	15
3.5	Machine learning, bias en berekening van fairness	16
3.6	Bias ontdekken met behulp van machine learning	17
3.7	Analyse van fairness	18
4	Conclusies, aanbevelingen en vervolgonderzoek	27
4.1	De belangrijkste uitkomsten van het fairness onderzoek	27
4.2	Toepassingen van dit onderzoek	29
4.3	Mogelijke tekortkomingen, risico's en aanbevelingen	30
4.4	Vervolgonderzoek	33
	Samenvatting	35
	Summary	39
	Dankwoord	43
	Referenties	45

II Appendices

A	Appendix A - Fairness analysesproces	51
A.1	Analyse van fairness	51
B	Appendix B - Fairness ratio's	61
B.1	Accuracy Equality	61
B.2	Equal Opportunity	62
B.3	Predictive Equality	64
B.4	Predictive Parity	65
B.5	Statistical Parity	66

LIJST VAN FIGUREN

Figuur 1.1	Un dimanche après-midi à l'Île de la Grande Jatte, George Seurat	1
Figuur 1.2	Onderzoekslijnen Learning Technology & Analytics	3
Figuur 1.3	Content en Context van Learning Technology & Analytics . . .	4
Figuur 3.1	Bias in een afbeelding van ChatGPT 40	13
Figuur 3.2	Identificatie van sensitieve kenmerken door een clusteranalyse	15
Figuur 3.3	Verdeling van de dataset in drie delen: de trainingsdata, de testdata en de validatiedata	19
Figuur 3.4	Toepassing van het model op een gemiddelde, fictieve student .	20
Figuur 3.5	Toepassing van het model op een gemiddelde, fictieve mannelijke student	21
Figuur 3.6	Ceteris Paribus Plot	21
Figuur 3.7	Partial Dependence Plot	22
Figuur 3.8	Distributie van de kans op retentie na 1 jaar per hoogste vooropleiding	23
Figuur 3.9	Fairness analyse per hoogste vooropleiding	24
Figuur 3.10	Fairness conclusies per variabele	24
Figuur 4.1	Fairness conclusies per variabele voor alle Ad en Bachelor opleidingen van De Haagse Hogeschool	29
Figuur A.1	Verdeling van de dataset in drie delen: de trainingsdata, de testdata en de validatiedata	53
Figuur A.2	Toepassing van het model op een gemiddelde, fictieve student .	54
Figuur A.3	Toepassing van het model op een gemiddelde, fictieve mannelijke student	55
Figuur A.4	Ceteris Paribus Plot	55
Figuur A.5	Partial Dependence Plot	56
Figuur A.6	Distributie van de kans op retentie na 1 jaar per hoogste vooropleiding	57
Figuur A.7	Fairness analyse per hoogste vooropleiding	58

Figuur A.8	Fairness conclusies per variabele	58
Figuur B.1	Confusion Matrix met 5 ratio's voor fairness	61

Deel I

INLEIDING – ONDERZOEK NAAR KANSENGELIJKHEID EN
FAIRNESS – CONCLUSIES

*We zien de dingen niet zoals ze zijn,
we zien de dingen zoals wij zijn.*

— Anais Nin

1

INLEIDING



Figuur 1.1: Un dimanche après-midi à l'Île de la Grande Jatte, George Seurat

Op het schilderij 'Un dimanche après-midi à l'Île de la Grande Jatte' zien we mensen die hun vrije zondagmiddag doorbrengen aan de oever van de Seine. De meeste mensen kijken naar de rivier. Eén meisje in een wit jurkje kijkt naar ons – de toeschouwers – terwijl ze bij haar moeder of oppas staat. Het tafereel kun je alleen zien door van een afstand te kijken. Dan nog blijft het beeld vaag. De kunstenaar, George Seurat, schilderde

duizenden stipjes om een geheel vormen. Deze techniek heet het *pointillisme*. Hij werkte er twee jaar aan, van 1884 tot 1886. Het beeld is precies, maar ook vluchtig: een indruk die snel kan vervagen. Het impressionisme, de kunststijl van Seurat, ontstond door de uitvinding van de verftube in 1841. Door de verftube konden kunstenaars langer buiten schilderen. In de buitenlucht ontdekten ze dat de werkelijkheid voortdurend verandert. Die verandering wilden ze vangen in hun schilderijen. Seurat kon dat als geen ander.

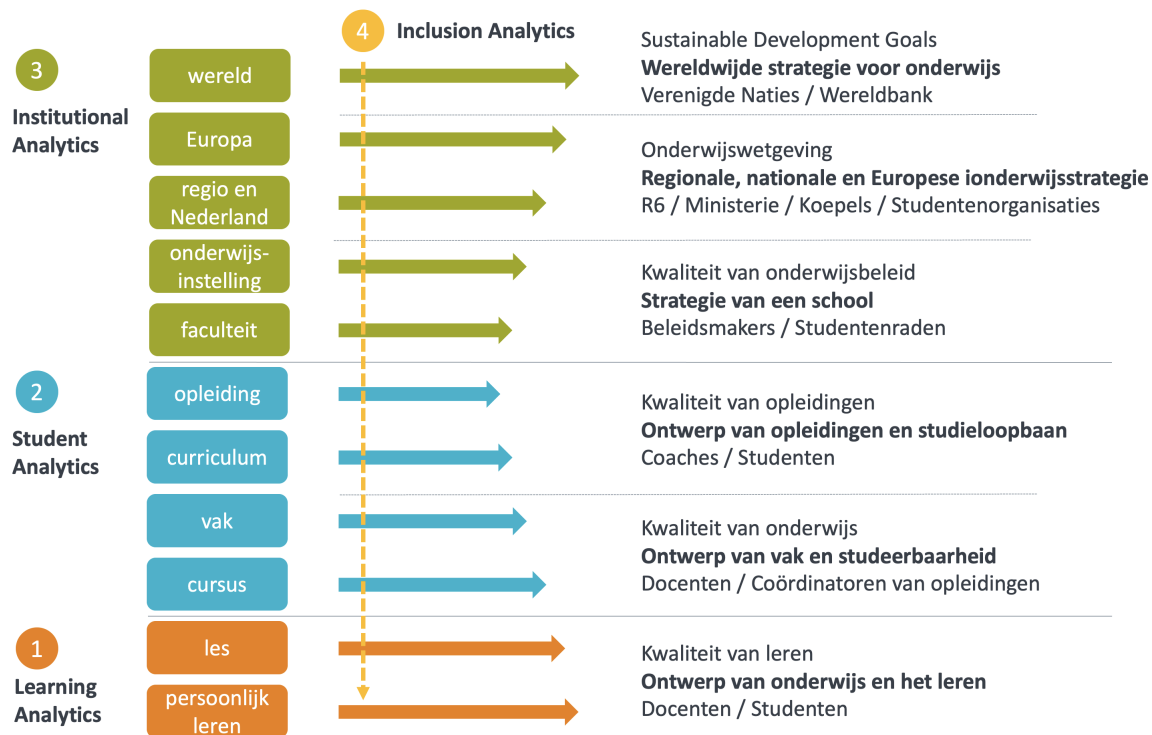
Dit schilderij hangt in Chicago en is te zien in de film 'Ferris Bueller's Day Off' uit 1986. Een film over orde en regels van het maatschappelijke systeem waarbinnen Ferris een dagje spijbelt. Hij sleurt daarin zijn vriend Cameron mee. Op dit punt in de film – het museumbezoek – verdrinkt Cameron in de ogen van het meisje in het schilderij. Hij lijkt te beseffen dat zijn leven weinig zin heeft als hij de regels van zijn vader blijft volgen.

Vandaag neem ik jullie mee op een reis waar ik twee jaar aan heb gewerkt. Een officiële spijbeldag waarop we afstand nemen en proberen een indruk te krijgen van patronen en regels in ons onderwijssysteem. Het onderwijs is ons canvas, de data over onze studenten zijn de verf, en de techniek die we gebruiken om van al die punten een beeld te maken is *learning analytics*.

1.1 LEARNING ANALYTICS, EDUCATIONAL DATA MINING EN SOCIALE RECHTVAARDIGHEID

Learning analytics is "het meten, verzamelen, analyseren en rapporteren van data van studenten en hun omgeving met als doel het leren en de leeromgeving te begrijpen en te optimaliseren" (Siemens & Long, 2011). Deze definitie klinkt breed, maar richt zich vooral op data die ontstaan door de wisselwerking tussen een student, medestudenten, de lesstof en de docent.

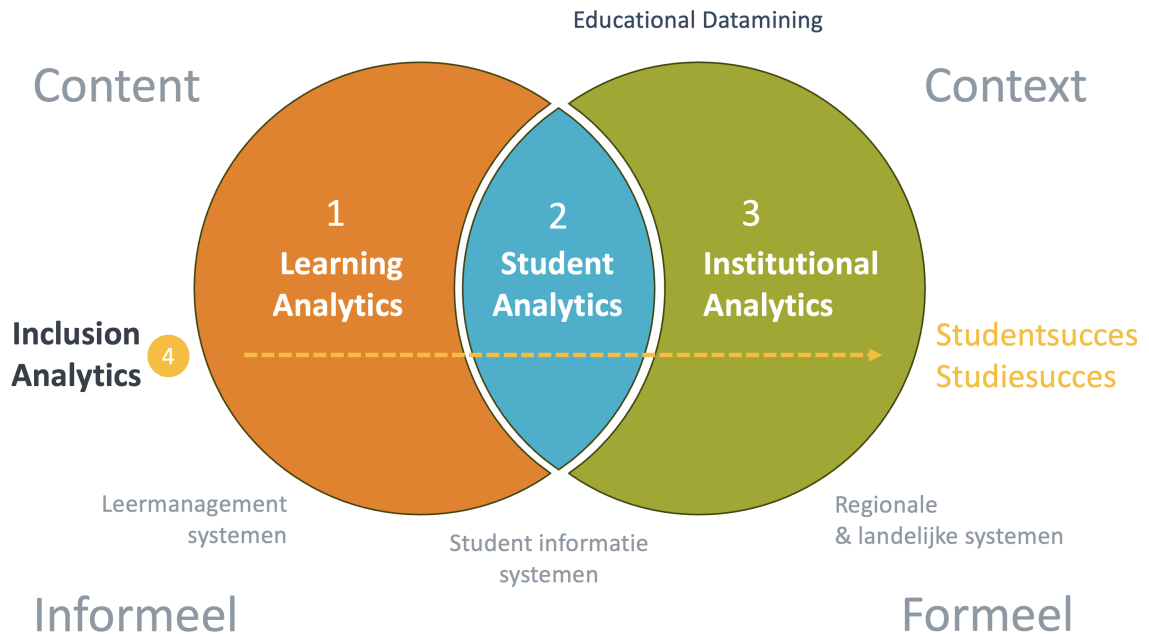
Dit is een van de onderzoekslijnen van het lectoraat, maar we onderzoeken meer (Figuur 1.2). Ons onderzoek richt zich ook op data op een hoger, meer systeemniveau. Denk aan vakken en vakcombinaties binnen opleidingen en studieloopbanen – dit noem ik *student analytics*. Nog hoger, op het niveau van de onderwijsinstelling, heet dit *institutional analytics*. In het vakgebied heten deze twee dimensies *Educational Data Mining* (EDM). Daarnaast heb ik een onderzoekslijn die dwars door deze verschillende vormen van analytics loopt, *inclusion analytics*, waarbij ik onderzoek in hoeverre er sprake is van kansengelijkheid in het onderwijs.



Figuur 1.2: Onderzoekslijnen Learning Technology & Analytics

Bij learning analytics staat de inhoud van het onderwijs centraal, de *content*; bij EDM gaat het om de *context* van het onderwijs. Omdat de data uit verschillende systemen komen, zijn deze gebieden vrij kunstmatig van elkaar gescheiden. Learning analytics gebruikt meer informele data uit leermanagementsystemen (LMS-en), terwijl EDM werkt met formele data uit administratieve systemen van onderwijsinstellingen, zoals het studentinformatiesysteem (SIS), Studielink en de Dienst Uitvoering Onderwijs (DUO) (Figuur 1.3).

Dit praktische verschil maakt het moeilijker om ons onderwijssysteem te begrijpen en het effect ervan op inclusie en succes van studenten te beoordelen. Als learning analytics zonder context wordt toegepast, kan dit tot het idee leiden dat een student vooral zichzelf moet aanpassen in houding of motivatie om succesvol te zijn in het onderwijs. Als de student hierin slaagt, is dat zijn, haar of hun eigen verdienste. De keerzijde is dat als dat niet lukt, dit aan hem-, haar- of henzelf te wijten is. In beide gevallen ligt het niet aan de context, het systeem. Dit idee, ook wel meritocratie genoemd, is wijdverspreid in het Nederlands onderwijs, maar legt volgens mij de verantwoordelijkheid voor het succes voor een groot deel op de verkeerde plek.



Figuur 1.3: Content en Context van Learning Technology & Analytics

Wie of wat is er dan wel verantwoordelijk voor succes? Voor succes in het hoger onderwijs worden de termen 'studiesucces' en 'studentsucces' gebruikt. De eerste term – studiesucces – richt zich op het belang van de opleiding en zou alleen kijken naar diplomarendement; de tweede – studentsucces – kijkt breder naar het welzijn van de student. Zelf geef ik toch de voorkeur aan de term studiesucces. Ik denk namelijk dat vrijwel elke student het hoger onderwijs wil verlaten met een diploma en dat de meeste studenten daarvan gelukkiger worden. Een tweede reden voor mijn voorkeur is, dat het de focus verlegt naar de studie en de organisatie daarvan. Het is de verantwoordelijkheid van iedereen in het hoger onderwijs om studenten daarin zo goed mogelijk te begeleiden en hen te behoeden voor problemen zoals depressie of *burn-out* op weg naar dit doel. Een goed voorbeeld van een interventie die hierbij past is de afschaffing van het Cum Laude diploma bij Geneeskunde studenten.

Als de verantwoordelijkheid ligt bij het onderwijssysteem, dan is het de vraag of het onderwijs dat we bieden wel sociaal rechtvaardig is voor alle studenten. Om deze vraag te beantwoorden, moeten we ons *bewust* zijn van mogelijke onrechtvaardigheid in het systeem dat wij hebben opgebouwd. Dit is het onderwerp van vandaag. *No Fairness without Awareness*.

1.2 DOEL EN OPBOUW INTREEREDE

In deze intreerede bespreek ik de uitkomsten van een statistisch onderzoek naar kansengelijkheid. Ik heb dit uitgevoerd voor alle bachelor- en associate degree-opleidingen aan onze hogeschool over de jaren 2011 tot en met 2022. Ik heb twee doelen. Ik wil allereerst de positie en het succes van minderheden binnen onze hogeschool verbeteren met een beter inzicht in kansengelijkheid. In de tweede plaats wil ik duidelijk maken hoe ik dit onderzoek heb uitgevoerd. Ik hoop dat ik daarmee dit soort onderzoek naar kansengelijkheid in het Nederlands onderwijs verder kan brengen.

Om dit te bereiken zal ik eerst dieper ingaan op het begrip kansengelijkheid en mijn persoonlijke motivatie met dit onderwerp. Vervolgens zal ik een onderzoeksmethode tonen waarmee we de mate van vooringenomenheid – ook wel *bias* genoemd – in studiedata kunnen analyseren. Hiermee wil ik een mate van eerlijkheid – *fairness* – kunnen vaststellen. Ik sluit af met conclusies, toepassingen van de inzichten, risico's en aanbevelingen, en plannen voor vervolgonderzoek.

Waar de behoeften van de wereld en je talenten elkaar kruisen,
daar ligt je roeping.

— Aristoteles

2

KANSENGELIJKHEID IN HET HOGER ONDERWIJS

In mijn onderzoek stel ik kansengelijkheid in onderwijs centraal. Mijn motivatie voor dit thema hangt samen met wie ik ben, met het gezin waar ik uit kom en hoe ik mijzelf door onderwijs heb kunnen emanciperen. Mijn ontwikkeling is beïnvloed door een zoektocht naar kansengelijkheid, zowel van mijzelf, als van mijn ouders. Ik wil dit persoonlijke verhaal vertellen, omdat ik denk dat vrijwel elke student zo'n verhaal heeft. Die verhalen probeer ik met behulp van *learning analytics* een gezicht te geven.

2.1 PERSOONLIJKE MOTIVATIE VOOR KANSENGELIJKHEID

Ik kom uit een onderwijsgezin – mijn ouders waren allebei docent. Zowel mijn vader als moeder komen uit een eenoudergezin. De ouders van mijn moeder waren in 1944 gescheiden, waardoor zij als enig kind werd opgevoed door haar moeder. Mijn vader, die iets ouder was, groeide op in een NSB-gezin en zijn vader ontnam zich aan het eind van de oorlog van het leven. Zij vonden elkaar in deze gemeenschappelijke, complexe achtergrond van eenzaamheid en maatschappelijke uitsluiting.

Hun opleiding tot docent was voor mijn ouders cruciaal. Onderwijs was de kans om een beter leven te leiden en sociaal en maatschappelijk erkend te worden. Als klein jochie van 3 jaar nam mijn moeder me al mee in haar klas, omdat ik niet bij de oppas wilde blijven. Het belang van onderwijs gaven mijn ouders aan ons mee. In ons gezin was ik de eerste die ging studeren in het hoger onderwijs: een eerstegeneratiestudent. Het was indirect ook een kans voor mijn ouders om hun ambities waar te maken. Voor mij was het een kans om mijzelf beter te begrijpen en uiteindelijk een *ander* leven te

leiden. Al jong wist ik dat ik homoseksueel was. Door een studie en nieuw leven in Utrecht kon ik het gereformeerde milieu waarin ik opgroeide achter me laten en – anders – toch een plaats geven. Uiteindelijk ben ik vandaag, hier op dit punt beland bij De Haagse Hogeschool. Nu lijkt het alsof ik dat aan mijzelf te danken heb. Dat is niet zo. Dat ik hier vandaag mag staan als lector Learning Technology & Analytics is een aaneenschakeling van kansen geweest die anderen – waarvan er velen vandaag aanwezig zijn – mij hebben gegeven. Die kansen heb ik kunnen oppakken doordat de randvoorwaarden voor mij gunstig waren. De eventuele *bottlenecks* die er waren, waren daardoor voor mij goed te doen. Dat samenspel van kansen, context en condities fascineert me.

2.2 HET BELANG VAN HOGER ONDERWIJS VOOR KANSENGELIJKHEID

Helaas werkt het Nederlands onderwijsstelsel ongelijkheid van kansen in de hand. Doordat leerlingen al op twaalfjarige leeftijd een keuze moeten maken voor een niveau, worden verschillen al vroeg uitvergroot (Copier, 2022). Eenmaal in een ‘lagere’ schoolloopbaan, kost het meer moeite om op te stromen naar het hoger onderwijs (Elffers, 2022). Toegang tot dat hoger onderwijs, wordt gezien als de ultieme vorm van kansengelijkheid. En dat is niet voor niets.

Instellingen voor hoger onderwijs spelen een belangrijke rol in kansengelijkheid, zoals benadrukt in de *Sustainable Development Goals* (SDG’s) van de Verenigde Naties. Doel 4.3 luidt: “Tegen 2030 gelijke toegang garanderen voor alle vrouwen en mannen tot betaalbaar en kwaliteitsvol technisch, beroeps- en hoger onderwijs, met inbegrip van de universiteit.” (“SDG Nederland - 4 Kwaliteitsonderwijs,” n.d.). Het hoger onderwijs is een motor voor persoonlijke ontwikkeling en versterkt de positie van jongeren in de samenleving (Baum, Ma, & Payea, 2010; Ma, Pender, & Welch, 2016; Perna, 2005; Tinto, 2012). Ik wil niet zeggen dat een mbo opleiding dit niet kan bieden, maar uit onderzoek blijkt dat hoogopgeleiden tevredener zijn in hun werk en leven, vaker een baan hebben, een hoger inkomen en een gezonder leven (Murray, 2009). Onze maatschappij belooft banen voor hoogopgeleiden beter, wat deze trend versterkt (Belfi et al., 2018).

2.3 BEPERKINGEN IN HET DENKEN OVER KANSENGELIJKHEID EN HOGER ONDERWIJS

De eerdere definitie uit de SDG's klinkt eerlijk, maar heeft twee beperkingen. Deze definitie gaat alleen over *toegang* tot het onderwijs en kansengelijkheid voor mannen en vrouwen. Dit is een begin, maar niet genoeg. Kansengelijkheid gaat niet alleen om gelijke toegang, maar ook om gelijke uitkomsten: gelijke kansen op een diploma en gelijke kansen op het vinden van een vervolgstudie of werk op niveau (Elffers, 2022). Om tot gelijke uitkomsten te komen, moeten we studenten soms juist ongelijk behandelen. Een tweede tekortkoming is dat de definitie zich beperkt tot kansengelijkheid op basis van het *geslacht*. Er zijn ook andere kenmerken die we moeten meenemen. Dit punt zal ik zo nader toelichten.

Ook de Vereniging Hogescholen en het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW) hanteren een te smalle kijk op kansengelijkheid. In 2019 heeft de VH speerpunten opgesteld voor het verbeteren van inclusiviteit en diversiteit in het hoger onderwijs (*Professionals voor morgen - Strategische agenda, Vereniging Hogescholen 2019-2023*, 2019). Daarin gebruikt de VH de term 'gelijke kansen' niet. Deze term lijkt voor de VH helaas voorbehouden voor de positionering van hogescholen zelf als volwaardige instituten naast universiteiten, zo blijkt uit het 'Position paper Gelijke kansen voor de hogeschoolstudent' uit 2023 (2023). Een beperkte visie op kansengelijkheid zien we ook terug in de recente uitgave van OCW over de langstudeermaatregel. Het rapport 'Talent op de juiste plek' (2024) gaat uit van de toegankelijkheid van het hoger onderwijs, maar neemt gelijke kansen op het behalen van een diploma niet mee. De voorgenomen langstudeermaatregel zal kansenongelijkheid verder in de hand werken. De fraaie titel van het rapport dekt dan ook zeker niet het effect dat we kunnen verwachten.

Over kansengelijkheid kunnen we vaststellen dat er verschillen in kansen zijn, maar dat het aanbieden van gelijke hulpmiddelen en begeleiding die niet oplossen. Ongelijk verdeelde hulpmiddelen en begeleiding kunnen leiden tot gelijke uitkomsten. Maar een wezenlijke systeemverandering is nodig om sociale rechtvaardigheid te bereiken. Daarvoor is het nodig om kansengelijkheid in het systeem te analyseren.

*Algorithms don't make things fair.
They repeat our past practices, our patterns.
They automate the status quo.
We have to check them for fairness.
Algorithms can be interrogated.
And they will tell us the truth every time.*

— C. O'Neil, 2017

3

ANALYSE VAN BIAS EN FAIRNESS IN STUDIEDATA

Nu we een beter beeld hebben van kansengelijkheid, wil ik ingaan op de begrippen *fairness* en *bias*¹. Ik koppel die aan de instroom, doorstroom en uitstroom van studenten in het hoger beroepsonderwijs.

3.1 DIVERSITEIT IN INSTROOM EN UITSTROOM

Niets blijkt zo divers als de instroom en de uitstroom in het hbo. Studenten komen binnen via mbo, havo, vwo, maar ook met een buitenlands diploma of via een 21+ toets. Zij schrijven zich in op voltijd studies, maar daarnaast hebben we deeltijd- en duale studies voor studenten die werken en studeren willen combineren. Studenten uit het vwo hebben vaak eerst een opleiding aan een universiteit geprobeerd, maar stappen dan alsnog over naar het hbo. En dan is er ook nog de instroom van andere opleidingen binnen het hbo, als studenten een opleiding voortijdig beëindigen. Gemiddeld blijkt in De Haagse Hogeschool maar liefst 1 op de 3 studenten een opleiding te starten als switcher ($M = 32,9\%$; interne switch: $M = 8,8\%$; externe switch: $M = 24,1\%$).

De uitstroom is minstens zo gevarieerd: natuurlijk vanwege een diploma, maar een groot deel van de studenten verlaat voortijdig hun opleiding, ongewenst of gewenst.

¹ Om verwarring te voorkomen gebruik ik in deze inleidende de oorspronkelijke Engelse termen in plaats van een Nederlandse vertaling.

Studenten vallen ongewenst uit doordat ze de studie of het studeren misschien te complex vonden, het een verkeerde studiekeuze was, of doordat zij al een diploma op zak hebben en daarmee een plan-B achter de hand. Gewenste redenen kunnen zijn om door te studeren aan een universiteit met een hbo-propedeusediploma of om te wisselen van opleidingsvorm. Gemiddeld valt aan De Haagse Hogeschool na 1 jaar 41% van de studenten uit ($M = 40,8\%$). Daarvan heeft gemiddeld 6% wel een propedeusediploma ($M = 6,3\%$).²

Wat we in ieder geval zien is dat studenten enorm creatief zijn in hun *student journey* en zo hun onderwijs flexibiliseren. Dit leidt tot de vraag: In welke mate is er sprake van kansengelijkheid in deze instroom en doorstroom voor verschillende groepen studenten? Een antwoord hierop gaan we geven op basis van een analyse van studiedata. Maar bevat die data geen bias? Juist het gegeven dat studiedata bias kan bevatten, gaat ons inzicht geven in de mate van kansengelijkheid.

3.2 ZORGEN OVER KANSENGELIJKHEID EN BIAS IN STUDIEDATA

Bij statistische analyses draait kansengelijkheid om fairness, bias en algoritmes (Barocas, Hardt, & Narayanan, 2023). *Bias* is “een systematische fout in besluitvormingsprocessen, die leidt tot oneerlijke uitkomsten” (Ferrara, 2023). Als bias negatieve effecten heeft op mensen, kan dit leiden tot discriminatie. Een *algoritme* is een formule waarmee een toekomstige uitkomst wordt voorspeld. *Fairness* houdt in dat algoritmes geen discriminerende of onrechtvaardige gevolgen mogen hebben (Shin, 2022). Een fictief voorbeeld van een algoritme is:

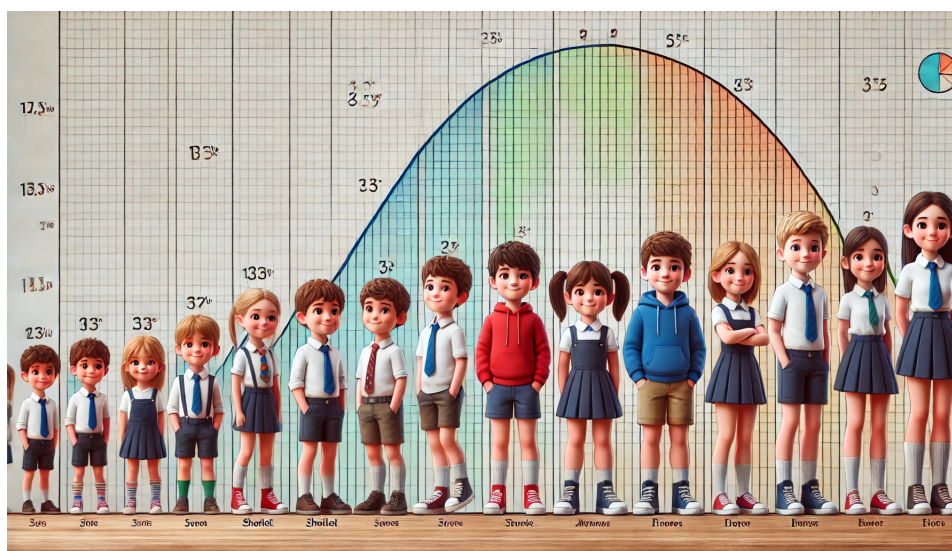
Geslacht = Vrouw + Leeftijd = 18 jaar + Aanmelddatum = 7 april + Vooropleiding = havo + Opleiding = B Communication & Multimedia Design: Kans op Retentie na 1 jaar = 62%, dus Retentie na 1 jaar = TRUE; immers, 62% kans rondt af naar 100%.

Ter illustratie van bias heb ik een voorbeeld toen ik ChatGPT-4 gebruikte met de DALL-E engine. Ik vroeg in het Engels een afbeelding te maken die – op basis van de lengte van leerlingen van een lagere school – een normale verdeling kon uitleggen³. De afbeelding die werd geproduceerd was op veel fronten problematisch;

² En dan is er nog de groep studenten die zich voor 1 februari tussentijds uitschrijft. Zij doen dit om te voorkomen een negatief BSA te krijgen. Vervolgens schrijven zij zich een jaar later weer in aan dezelfde opleiding om die alsnog af te maken. Het is geen uitval, maar het is ook niet echt doorstroom. Gemiddeld is dit slechts 1% ($M = 0,9\%$) van de herinschrijvers.

³ Uitgevoerd in ChatGPT-4o op 27-10-2024. De oorspronkelijke prompt luidde: “Imagine you work at a primary school with 200 children aged 4 to 12. If you ask the children to line up according to height, with shorter children on one side and taller children on the other, you would observe a pattern. Most children would have average heights for their age, so they would cluster in the middle. However, a few

voor deze introereede is het probleem de weergave van een groep witte kinderen in een Amerikaanse kostschool outfit (Figuur 3.1). Blijkbaar is dit het gemiddelde beeld van blijde leerlingen op de basisschool.



Figuur 3.1: Bias in een afbeelding van ChatGPT 4o

Vanuit meerdere kanten zijn er zorgen geuit over de toepassing van algoritmes in het onderwijs en kansengelijkheid. Een inmiddels bekend voorbeeld is het algoritme waarmee DUO eventuele fraude met uitwonende beurzen opspoorde (Ersoy & Gaag, 2023). Dit algoritme bestempelde studenten met een niet-Europese migratieachtergrond opvallend vaak als frauderend. Bij nader onderzoek bleek dat het gegeven of een student in de buurt van zijn, haar of hun ouders woonde een belangrijk selectiekenmerk was in het algoritme. Dit komt bij studenten met een migratieachtergrond vaker voor (“Vervolgonderzoek bevestigt indirecte discriminatie controles uitwonendenbeurs,” 2024). Sinds vorige week is bekend dat deze studenten hiervoor gecompenseerd zullen worden (Stift, 2024).

Het College van de Rechten van de Mens heeft onderzoek laten doen om risico’s op discriminatie, uitsluiting en kansenongelijkheid in kaart te brengen (Smeets, Geurts, & Helvoirt, 2024; “ScienceGuide,” 2024). Ook de Autoriteit Persoonsgegevens (AP) ziet grote uitdagingen als het gaat om kansengelijkheid en het gebruik van algoritmes

children would be much shorter or much taller than the average, standing at the ends of the line. This arrangement creates a curve when you plot the heights on a graph: the highest point is in the middle, where most children are, and it gradually slopes down on both sides, where the fewer, very short and very tall children stand. This shape is known as a bell curve or normal distribution, which is common in natural measurements like height, weight, and test scores.”

(Sectorbeeld Onderwijs 2021-2023, 2024). Met name is er zorg dat bias in de studiedata die worden gebruikt voor algoritmes, leidt tot kansenongelijkheid en discriminatie. Het valt op dat er in dit onderzoek geen concrete voorbeelden zijn gevonden in de praktijk. Toch wordt het advies gegeven dat docenten nagaan of er mogelijk sprake is van benadeling in de uitkomsten. Tot slot is er de Europese AI act die een aantal toepassingen van algoritmes in het onderwijs als hoog risico heeft aangemerkt. Onderwijsinstellingen zijn verplicht aan te tonen dat er geen sprake is van bias in de algoritmes die zij toepassen of gebruiken via software (Engelfriet, 2024).

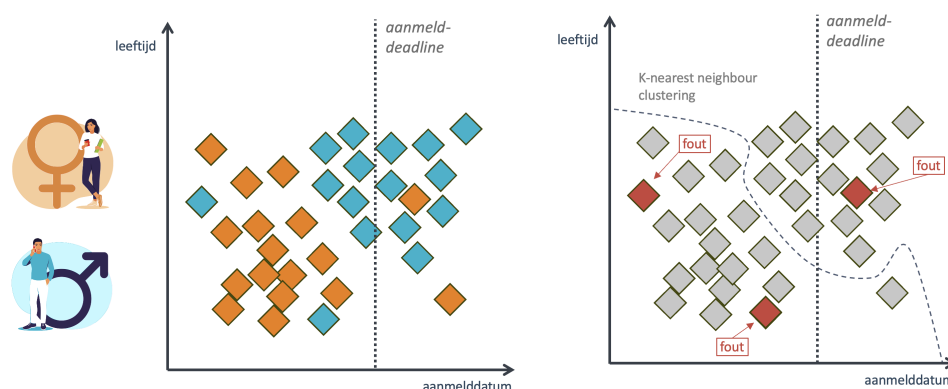
Er is dus een groeiende maatschappelijke zorg over mogelijke kansenongelijkheid en discriminatie in het onderwijs, maar de verantwoordelijkheid voor een oplossing wordt vaak gelegd bij onderwijsinstellingen en in het bijzonder docenten. Dit is naar mijn mening een bijna onmogelijke opdracht. Met name omdat wij onderdeel uitmaken van een *systeem* dat ongelijkheid in de hand werkt. Daardoor zit kansenongelijkheid ook in studiedata, in algoritmes, in AI en alle beslissingen die daar verder uit worden afgeleid. We moeten dan ook bij studiedata zelf beginnen om te analyseren welke groepen stelselmatig minder kansen hebben, zodat we daar bewust op kunnen handelen.

3.3 SENSITIEVE KENMERKEN VERSUS BIJZONDERE PERSOONSgegevens

Als we fairness onderzoeken, kijken we naar *sensitieve kenmerken*, niet te verwarren met *bijzondere persoonsgegevens*. Het zijn gegevens zoals etnische achtergrond en gezondheid. Sensitieve kenmerken zijn gegevens zoals genderidentiteit, handicap, leeftijd, sociale afkomst, geboorte en eigendom, taal en politieke overtuigingen uit het Handboek over het Europese non-discriminatierecht (*European Union Agency for Fundamental Rights*, 2018). Bijzondere persoonsgegevens zijn bepaald in de Algemene verordening gegevensbescherming (AVG). Deze kenmerken overlappen slechts voor een deel met elkaar. Sensitieve persoonsgegevens zijn kenmerken waarop gediscrimineerd kan worden. Bijzondere persoonsgegevens worden beschermd om schending van privacy te voorkomen

Het betrekken van sensitieve kenmerken in analyses is niet onomstreden. Ook al zijn het geen bijzondere persoonsgegevens, volgens sommigen is het toch beter deze gegevens niet op te nemen in analyses om bias te voorkomen (Pedreshi, Ruggieri, & Turini, 2008). Dit is een misvatting (Hardt, Price, & Srebro, 2016): het weglaten van deze kenmerken betekent niet dat ze geen invloed hebben op de resultaten van een analyse of beslissingen die daaruit voortkomen. Via andere variabelen in een dataset kunnen ze toch tot uiting komen. Stel we hebben een dataset met de datum van

aanmelding voor een studie, de leeftijd en het geslacht van studenten. Als we dit analyseren met een clusteranalyse, kan een algoritme op basis van patronen in leeftijd en aanmelddatum alsnog een verschil tussen mannen en vrouwen zien (Figuur 3.2). We willen mogelijke bias in bestaande data juist ontdekken. Zo kunnen we *voorkomen* dat bias bij de ontwikkelingen van algoritmen in het hoger onderwijs in beleid wordt omgezet en leidt tot institutionele discriminatie (Pedreshi, Ruggieri, & Turini, 2008). Ook hier geldt *No fairness without awareness* (Dwork et al., 2011). Om kansengelijkheid beter te begrijpen en te bevorderen, moeten we deze gegevens niet weglaten, maar juist wel analyseren.



Figuur 3.2: Identificatie van sensitieve kenmerken door een clusteranalyse

3.4 HUIDIG ONDERZOEK NAAR FAIRNESS IN HET (HOGER) ONDERWIJS

Er is wel een belangrijke kanttekening. Onderzoek naar fairness en bias is voornamelijk afkomstig uit de Verenigde Staten (VS). De datasets die van belang waren bij de ontwikkeling van dit type onderzoek, waren datasets voor het toekennen van leningen, van een borgsom na een arrestatie, van een baan of – relevant voor ons werkveld – van een plek in een universitaire opleiding (Dwork et al., 2011). De hoofdvraag in deze onderzoeken is of etnische minderheden gelijk werden behandeld in deze procedures. Dit geeft – bijzonder genoeg – in het onderzoek zelf een bias. Zoals we al eerder zagen is de Amerikaans context is niet per se representatief voor de Europese of Nederlands context. Er zijn biases naar sensitieve kenmerken die nationaal of regionaal formeel onderkend zijn, zoals 'veteraan' of 'etniciteit' in de VS. De meest onderzochte kenmerken in fairness in het hoger onderwijs zijn etniciteit, nationaliteit en geslacht (Baker & Hawn, 2022). Andere relevante kenmerken, zoals taal, functiebeperking,

herkomst uit de stad of van het platteland (*urbanicity*), sociaal economische status en vooropleiding zijn echter onderbelicht.

3.5 MACHINE LEARNING, BIAS EN BEREKENING VAN FAIRNESS

We zijn nu aanbeland bij de methode van het fairness onderzoek. Om fairness te onderzoeken gebruiken we algoritmes die met behulp van machine learning zijn ontwikkeld. Machine learning is een deelgebied van *Artificial Intelligence*. Het analyseert patronen in historische gegevens en past die vervolgens toe op nieuwe gegevens als voorspelling in de vorm van een formule, het algoritme. De mate waarin een algoritme een nieuwe, nog onbekende situatie goed voorspelt noemen we de *accuraatheid* van het model.

Bias in algoritmes houdt in dat de kansen in de uitkomsten van de voorspelling ongelijk zijn verdeeld tussen groepen. Dit zal vrijwel altijd zo zijn. Het kan bijvoorbeeld betekenen dat in een selectieprocedure vwo studenten een groter aandeel hebben in het aantal studenten dat is toegelaten. Laten we dit concreet maken. Stel er is een selectieve opleiding met 100 plekken. Daarvoor melden zich 225 studenten aan: 150 vrouwen en 75 mannen. Bij de selectie wordt er gekeken naar eerdere cijfers op de middelbare school, ervaring in de zorg, een motivatiebrief en een gesprek. Van de vrouwen worden er 75 geselecteerd, van mannen 25. Is dit een eerlijke selectie?

Of het oneerlijk is – een gebrek aan fairness – hangt af van een criterium dat we daarvoor hanteren. In de VS is voor dit soort verschillen een wettelijke maatstaf ontwikkeld om te bepalen of een procedure nog wel eerlijk is: het fairness criterium van statistische gelijkheid (“Code of Federal Regulations. Section 4d, uniform guidelines on employee selection procedures (1978),” 1978). Dit criterium houdt in dat in een selectieprocedure de tweede groep — de beschermde groep — niet kleiner mag zijn dan $\frac{4}{5}$ deel van de eerste groep — de bevoorrechte groep. We gaan terug naar het voorbeeld. De verhouding geselecteerd/niet geselecteerd per groep is respectievelijk 50% en 33%. Volgens het $\frac{4}{5}$ criterium is er sprake van oneerlijkheid: er zijn onevenredig minder mannen die worden geselecteerd. Een eerlijke verdeling zou zijn: 75 vrouwen en 30 mannen (40 vrouwen = 50% van 150 studenten; $\frac{4}{5} * 50\% = 40\%$; $40\% * 75$ mannen = 40 studenten). Een eerlijke verdeling voor 100 plekken in instroom met $\frac{4}{5}$ verhoudingen zou zijn: 70 vrouwen en 30 mannen. Met deze $\frac{4}{5}$ formule is bias in selectieprocedures eenvoudig door te rekenen, maar voor analyses op eerlijke doorstroom is er meer nodig.

3.6 BIAS ONTDEKKEN MET BEHULP VAN MACHINE LEARNING

Zoals ik al heb toegelicht, wordt machine learning gebruikt om voorspellingen te maken om deze in de onderwijspraktijk toe te passen in besluitvorming. Een tweede toepassing is het doen van wetenschappelijk onderzoek (Shmueli, 2010; Shmueli & Koppius, 2011); daarbij gebruiken we voorspelmodellen om structuren in data beter te begrijpen. Beide vormen van machine learning passen uitstekend bij het doen van toegepast onderzoek in het hoger beroepsonderwijs. Een belangrijk verschil met alleen statistische verbanden, zoals het verband tussen twee variabelen (de correlatie), is dat een voorspelmodel via de mate van accuraatheid ook uitdrukt hoe toepasbaar inzichten kunnen zijn in de praktijk. Laten we hier op inzoomen, want dit is van belang voor het ontdekken van bias.

De methode die we hanteren voor ons onderzoek, maakt gebruik van het gegeven dat algoritmes bestaande patronen vinden voor zo goed mogelijke voorspellingen. Een algoritme heeft daarbij geen ethisch kompas. Het kijkt wel naar meerderheden of minderheden, maar pleegt geen ethische interventies. Het maakt wel onderscheid, maar discrimineert niet. Het neemt die combinatie van factoren die voor de uitkomst die wordt onderzocht het beste werkt – in ons onderzoek ‘retentie na 1 jaar’. Als er bias is in de onderliggende data, zal die door het algoritme zonder oordeel worden omgezet naar een formule, het algoritme. Bijvoorbeeld dat in een bepaalde opleiding een man minder kans heeft om na 1 jaar door te studeren dan een vrouw. We onderzoeken de verschillen in de kwaliteit van de voorspelde kansen.

Maar, hoe komen we aan het punt waarbij dat hier geen sprake is van een normale dynamiek onder studenten? Want er kunnen allerlei ‘goede’ redenen zijn waarom een student niet doorstudeert na het eerste studiejaar. We lossen dit vraagstuk op door in de kwaliteit van de voorspelling te onderzoeken of de verschillen die we daar vinden tussen groepen onevenredig is volgens het 4/5 criterium.

3.7 ANALYSE VAN FAIRNESS

Het proces om fairness te berekenen kent drie stappen: 1) het ontwikkelen van voorspelmodellen op basis van historische gegevens⁴, 2) het bepalen van het belang van variabelen en 3) het berekenen van fairness⁵.

Ik vat stappen 1 en 2 kort samen. In een bijlage van deze introereede heb ik deze verder uitgewerkt (zie Appendix A). Bij stap 3, de fairness analyse, sta ik iets langer stil.

3.7.1 *Stap 1 - Het ontwikkelen van voorspelmodellen*

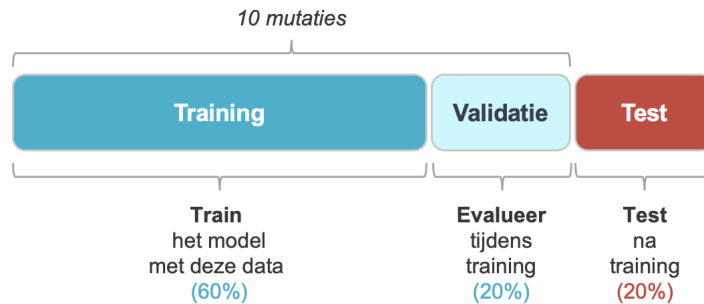
De eerste stap is het prepareren van de data en het bouwen van een aantal voorspelmodellen om Retentie na 1 jaar te voorspellen per opleiding.

- We verzamelen historische studiedata per student per studiejaar en schonen die op (*datapreparatie*). Ontbrekende gegevens vullen we zoveel mogelijk op (*imputatie*).
- We onderzoeken welke variabelen sterk met elkaar samenhangen en maken daar eventueel een selectie uit (*multicollineariteit*).
- Vervolgens splitsen we de dataset in een training-, test- en validatieset, zodat de prestaties van de modellen nauwkeurig kunnen worden beoordeeld (datapartities; Figuur A.1).
- We zorgen ervoor dat er een gelijkmatige verdeling is van de uitkomstmaat per set (stratificatie).
- Hierna maken we meerdere voorspelmodellen (logistische regressie en random forest). We kiezen modellen waarvan goed is uit te leggen welke variabelen er onderdeel van uitmaken en wat de bijdrage per variabele is aan de voorspelling.
- Tot slot bepalen we per model hoe goed het presteert: in welke mate een voorspelmodel in staat is om positieve uitkomsten als positief te voorspellen en negatieve uitkomsten als negatief. Dit noemen we de *accuraatheid* van het model.

⁴ Voor dit onderzoek hebben we gebruik gemaakt van het tidymodels package in R (Kuhn & Wickham, 2020).

⁵ Voor het onderzoek naar variabelen en fairness hebben we gebruik gemaakt van de DALEX, DALEXtra en fairmodels packages in R (Biecek, 2018; Maksymiuk, Gosiewska, & Biecek, 2020; Wiśniewski & Biecek, 2022).

In ons voorbeeld betekent dit dat het model goed voorspelt welke student er wel of niet doorstudeert na 1 jaar. Het model met de beste prestaties vormt de basis voor het vervolg.



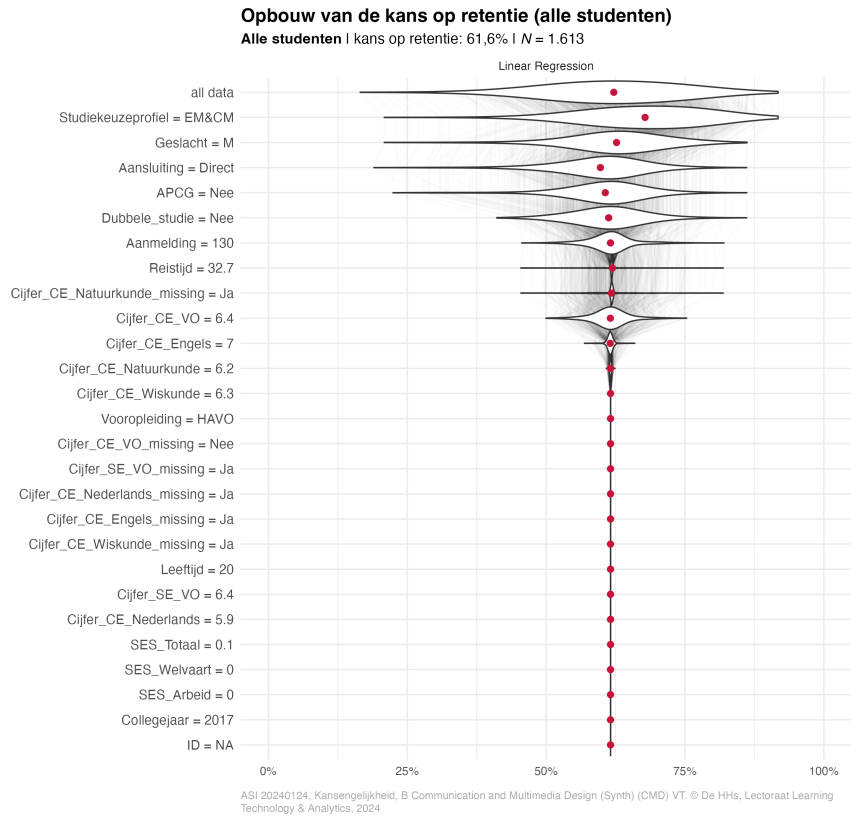
© De Haagse Hogeschool, Lectoraat Learning Technology & Analytics, 2024

Figuur 3.3: Verdeling van de dataset in drie delen: de trainingsdata, de testdata en de validatiedata

3.7.2 Stap 2 - Het bepalen van het belang van variabelen

Het tweede deel van de analyse bestaat uit een beter begrip van de variabelen in het voorspelmodel. We gaan per opleiding verder met het model met de hoogste accuraatheid om het belang van variabelen in te schatten voor verschillende groepen studenten.

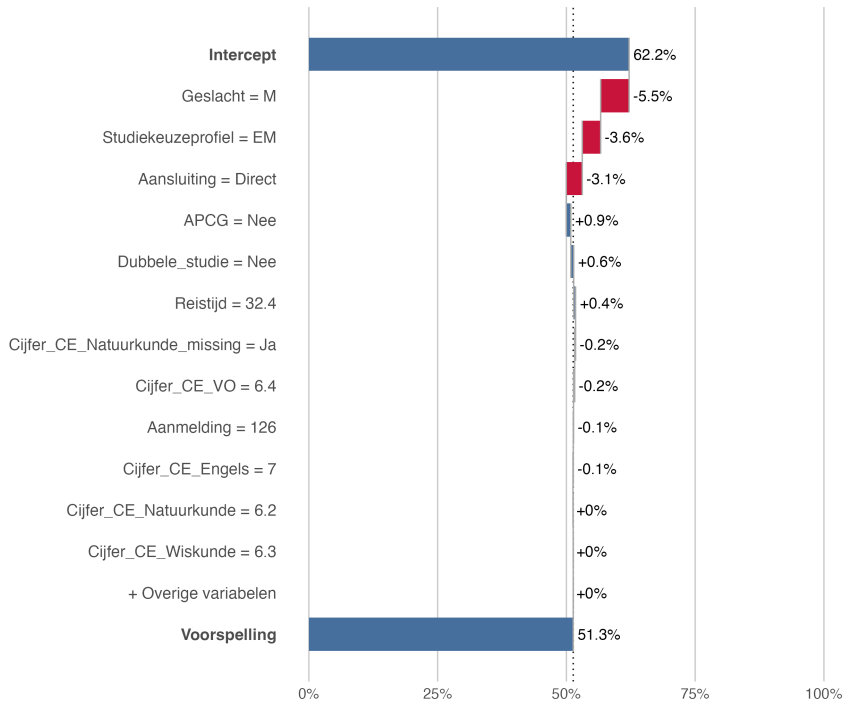
- We passen het model toe op een 'gemiddelde student' binnen de opleiding, die we berekenen op basis van gemiddelden (voor continue variabelen) en meest voorkomende waarden (voor categorische variabelen).
- Dit herhalen we voor verschillende groepen (bijv. geslacht, vooropleiding) door per groep een gemiddelde student te berekenen van die groep.
- We passen daarop extra analyses toe – zoals Shapley, Ceteris Paribus en Partial Dependency – om de volgorde van variabelen, de effecten van variabelen en de variatie binnen groepen beter te begrijpen. Zie Figuur A.2, Figuur A.3, Figuur A.4 en Figuur A.5.



Figuur 3.4: Toepassing van het model op een gemiddelde, fictieve student

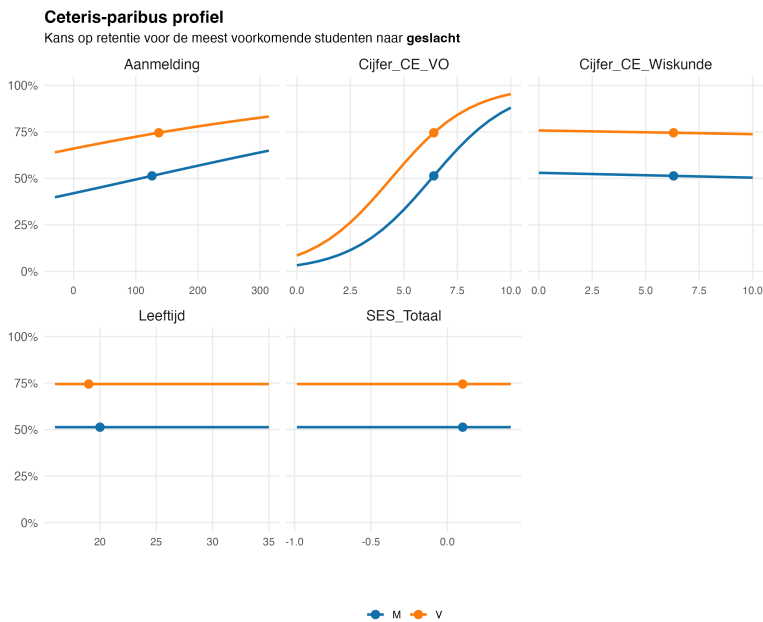
Opbouw van de kans op retentie naar geslacht

Geslacht: M | kans op retentie: 51,3% | N = 915 van 1.613 (56,7%)



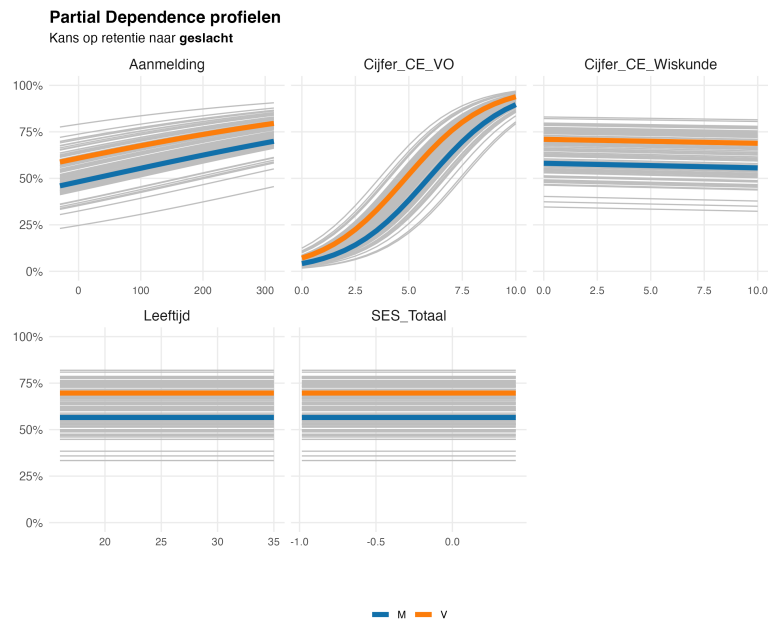
ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur 3.5: Toepassing van het model op een gemiddelde, fictieve mannelijke student



ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur 3.6: Ceteris Paribus Plot



ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur 3.7: Partial Dependence Plot

3.7.3 Stap 3 - Het berekenen van fairness

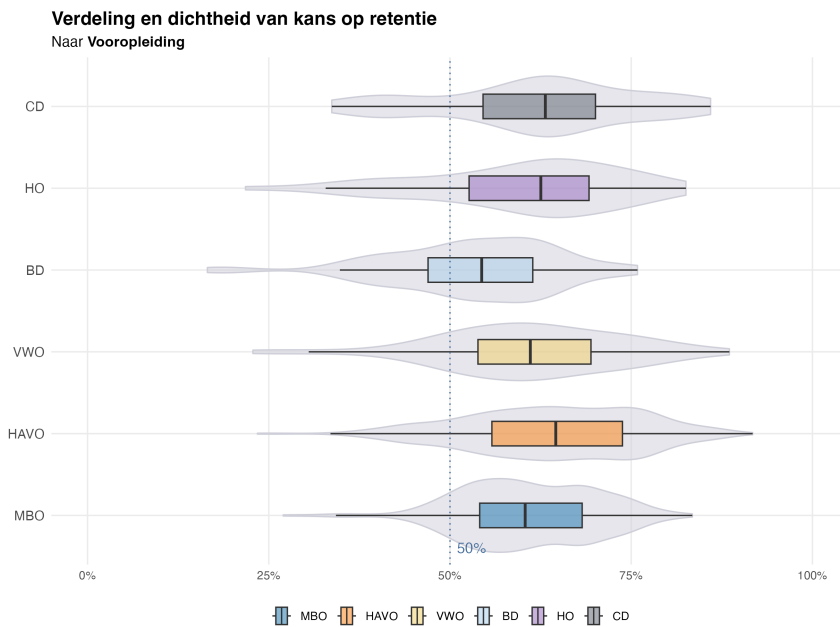
Het derde en laatste deel van het proces is de berekening van fairness:

- Eerst analyseren we de kansverdeling binnen groepen; hiermee krijgen we een goed beeld per groep wat de verdeling is van de kansen in het voorspelmodel en kunnen die onderling ook goed vergelijken. Zie Figuur A.6.
- Vervolgens passen we het 4/5 criterium toe op een aantal prestatiewaarden van het model, waaronder accuraatheid⁶. Als bevoorrechte groep hanteren we de meest voorkomende subgroep binnen een variabele⁷. Als er twee of meer waarden buiten de bandbreedte van het 4/5 criterium vallen is er sprake van een gebrek aan fairness voor een beschermde groep ten opzichte van de bevoorrechte groep. Dit kan negatief, maar ook juist positief uitvallen voor sommige groepen. Zie Figuur A.7.

⁶ In totaal zijn 5 ratio's beoordeeld: Accuracy Equality, Equal Opportunity, Predictive Equality, Predictive Parity en Statistical Parity. Zie Appendix B voor een toelichting op deze ratio's.

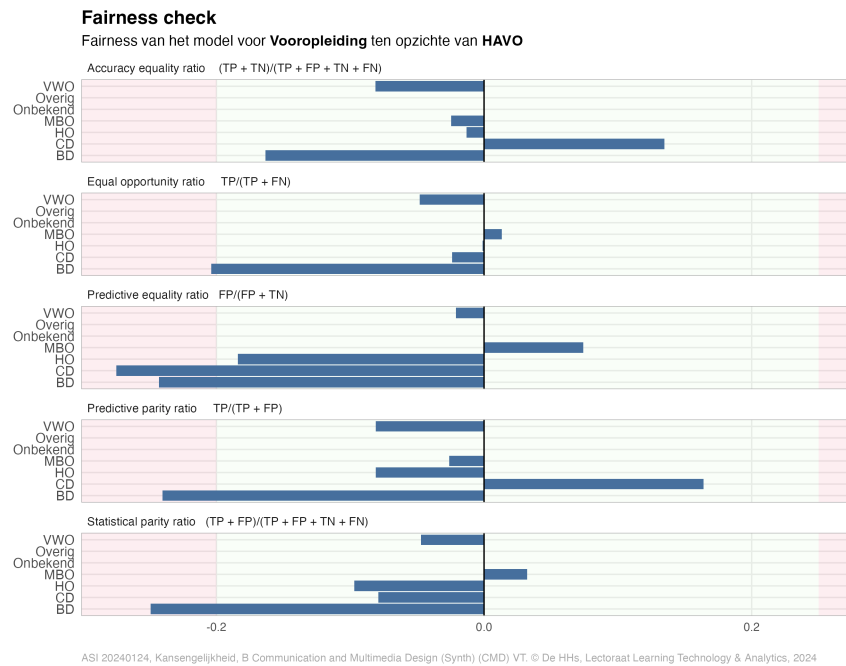
⁷ Stel dat er meer mannen dan vrouwen zijn in een opleiding, dan is Man de bevoorrechte groep; bij een gelijke stand wordt een van de twee gekozen.

- Tot slot vatten we de bevindingen samen in een tabel om te zien waar er bias is en voor welke groepen. Daarbij maken we een onderscheid tussen groepen waarvan we voldoende data hebben om een uitspraak te doen of niet⁸. Zie Figuur A.8.



Figuur 3.8: Distributie van de kans op retentie na 1 jaar per hoogste vooropleiding

⁸ NB. Het kan zijn dat de bevoorrechte groep juist de groep is die last heeft van bias. Dit is vast te stellen als alle andere groepen een negatieve of positieve bias hebben



Figuur 3.9: Fairness analyse per hoogste vooropleiding

Variabele	Groep	Bias	Geen Bias	Negatieve Bias	Positieve Bias
<i>Geslacht</i>	V	Nee	5	0	0
<i>Vooropleiding</i>	BD	Ja	1	3	1
	CD	Ja	0	0	5
	HO	Nee	5	0	0
	MBO	Ja	2	3	0
	VWO	Ja	3	0	2
<i>Aansluiting</i>	2e Studie	Ja	2	2	1
	Na CD	Ja	0	0	5
	Switch extern	Nee	4	0	1
	Switch intern	Nee	4	1	0
	Tussenjaar	Ja	2	0	3

Figuur 3.10: Fairness conclusies per variabele

Deze onderzoeksprocedure hebben we doorlopen voor alle associate degree en bachelor-opleidingen van De Haagse Hogeschool in elke variant. Per opleiding hebben

we de uitkomsten opgeschreven in een zichzelf uitleggende website. Vanaf vandaag is elke analyse beschikbaar voor elke opleiding.

*In God we trust.
All others must bring data.*
— W. Edwards Demming

4

CONCLUSIES, AANBEVELINGEN EN VERVOLGONDERZOEK

Nu we alle stappen doorlopen hebben om fairness te analyseren, ben ik toegekomen aan de bespreking van de belangrijkste uitkomsten, de mogelijke toepasbaarheid van deze aanpak, eventuele tekortkomingen en aanbevelingen, en het vervolgonderzoek dat we willen gaan uitvoeren binnen het lectoraat.

4.1 DE BELANGRIJKSTE UITKOMSTEN VAN HET FAIRNESS ONDERZOEK

- **Fairness is te analyseren en te bespreken** – De eerste uitkomst is dat deze aanpak een herhaalbare en schaalbare onderzoeksmethode blijkt te zijn om bias te bepalen in studiedata. Het onderzoek blijkt voor alle opleidingen aan onze hogeschool uit te voeren en fairness kan met vrij basale gegevens. Uit de eerste besprekingen met collega's van opleidingen blijkt de aanpak goed uit te leggen en direct inzicht te geven in groepen studenten die onevenredig vaker niet doorstromen na hun eerste studiejaar.
- **Er is nog wel veel te winnen aan het verbeteren van fairness** – De tweede uitkomst is dat bij de meeste van de 66 opleidingen er sprake is van een gebrek aan fairness voor sommige groepen studenten (zie [Figuur 4.1](#)). Slechts 6 opleidingen vormden hierop een uitzondering, waarvan 1 een selectieve opleiding was. Verder is het opvallend dat bij slechts 1 opleiding er een negatieve bias is naar retentie na 1 jaar voor vrouwen en bij 20 opleidingen een negatieve bias naar mannen. Bij 20 opleidingen is er op kansgelijkheid voor mbo studenten nog veel te winnen, terwijl in 29 opleidingen vwo studenten juist onevenredig veel kans hebben.

- **Sensitieve kenmerken komen tot uitdrukking in niet-sensitieve kenmerken**
 - Een derde uitkomst is dat sommige variabelen onbelangrijk lijken, maar een achterliggend kenmerk uitdrukken. Het gegeven gemiddeld cijfer Centraal Examen, bijvoorbeeld, is relevant omdat dit ontbreekt voor mbo-studenten en studenten met een buitenlandse vooropleiding. Dat maakt deze variabele waardevoller in voorspelmodellen dan de achterliggende variabele, de hoogste vooropleiding. Variabelen als hoogste vooropleiding, middelbare schoolcijfers en inschrijfdatum zijn vaak proxies voor sociaal-economische status en mogelijk migratieachtergrond. Ze lijken capaciteiten of motivatie weer te geven, maar worden beïnvloed door achterliggende sociaal-economische factoren. Meer onderzoek hiernaar is nodig.
- **Er is potentie voor gebruik van deze inzichten, maar nog weinig vraag** – Een vierde uitkomst betreft het gebruik van deze analyses. Hoewel het instellingsplan van De Haagse Hogeschool inzet op het wegnemen van studieknelpunten “Onderzoekend leren met impact,” 2022, kwam het verzoek om kansengelijkheid te onderzoeken nog niet vaak vanuit de onderzochte opleidingen. Alleen binnen de sector Economie zijn afspraken gemaakt om kansengelijkheid te beoordelen (*Taakopdracht sectorale verkenning hoger economisch onderwijs (HEO) 2020-2021, 2020*). Meestal richten verzoeken zich op het verhogen van instroom of verlagen van uitval. Het goede nieuws is: verbetering van kansengelijkheid kan instroom verhogen en uitval verlagen. Ook hier geldt *No fairness without awareness*.

Faculteit	Opleiding	Geslacht		Vooropleiding						Aansluiting					
		M	V	MBO	HAVO	VWO	BD	HO	CD	Direct	Tussenjaar	Switch intern	Switch extern	Na CD	2e Studie
		Positief	Negatief	NTB	Positief	Onbeslist	Positief	Positief	Onbeslist	Negatief	Nee	Nee	NTB	Onbeslist	Negatief
		Nee	Nee	Negatief	NTB	Positief	Negatief	Negatief	Positief	NTB	Nee	Negatief	Nee	Onbeslist	Nee
		Nee	Nee	NTB	Positief	Onbeslist	Onbeslist	Positief	Positief	Nee	Nee	Nee	NTB	Positief	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Negatief	Nee	Negatief	NTB	Negatief	Nee	Negatief	Negatief	Negatief
		Nee	Nee	NTB	Positief	Onbeslist	Nee	Positief	Positief	Negatief	Nee	Negatief	NTB	Onbeslist	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Nee	Negatief	Negatief	NTB	Nee	Nee	Nee	Negatief	Negatief
		Nee	Nee	Nee	Onbeslist	NTB	Onbeslist	Onbeslist	Onbeslist	NTB	Nee	Negatief	Negatief	Geen data	Onbeslist
		Nee	Nee	Geen data	Onbeslist	Positief	NTB	Onbeslist	Onbeslist	NTB	Nee	Onbeslist	Negatief	Geen data	Onbeslist
		Nee	Nee	Negatief	Nee	Nee	NTB	Negatief	Nee	NTB	Nee	Nee	Nee	Nee	Negatief
		Nee	Nee	Geen data	Onbeslist	Nee	NTB	Onbeslist	Onbeslist	NTB	Nee	Onbeslist	Nee	Onbeslist	Onbeslist
		Negatief	Positief	Nee	NTB	Positief	Nee	Negatief	Nee	NTB	Nee	Nee	Nee	Onbeslist	Nee
		Nee	Nee	Negatief	NTB	Positief	Onbeslist	Nee	Positief	NTB	Positief	Nee	Nee	Positief	Onbeslist
		Negatief	Positief	Nee	NTB	Positief	Nee	Negatief	Negatief	NTB	Nee	Nee	Negatief	Negatief	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Positief	Positief	Nee	Nee	Onbeslist	NTB	Positief	Geen data
		Nee	Nee	Nee	NTB	Nee	Nee	Nee	Nee	NTB	Nee	Negatief	Negatief	Nee	Negatief
		Nee	Nee	Negatief	Negatief	Negatief	NTB	Negatief	Nee	NTB	Nee	Negatief	Negatief	Nee	Negatief
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Onbeslist	Positief	Negatief	NTB	Onbeslist	Nee	Nee	Geen data
		Negatief	Positief	Nee	NTB	Positief	Onbeslist	Positief	Positief	NTB	Positief	Nee	Positief	Positief	Onbeslist
		Nee	Nee	Negatief	Nee	Nee	NTB	Onbeslist	Nee	NTB	Positief	Nee	Nee	Positief	Negatief
		Onbeslist	Nee	NTB	Onbeslist	Onbeslist	Onbeslist	Nee	Onbeslist	Onbeslist	Nee	Onbeslist	NTB	Onbeslist	Geen data
		Nee	Nee	Nee	NTB	Nee	Nee	Nee	Nee	NTB	Nee	Nee	Nee	Nee	Onbeslist
		Nee	Nee	Negatief	NTB	Positief	Onbeslist	Nee	Onbeslist	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Nee	Nee	NTB	Nee	Nee	Negatief	Negatief	Nee	Nee	Nee	Nee	Nee	Onbeslist
		Nee	Nee	NTB	Negatief	Onbeslist	Positief	Positief	Onbeslist	Nee	Nee	Onbeslist	NTB	Onbeslist	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Nee	Onbeslist	Negatief	Nee	Onbeslist	NTB	Onbeslist	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Nee	Nee	Nee	Negatief	Nee	Negatief	Nee	Positief	Negatief
		Negatief	Positief	Nee	NTB	Positief	Negatief	Negatief	Onbeslist	NTB	Nee	Negatief	Nee	Onbeslist	Onbeslist
		Nee	Nee	Geen data	Onbeslist	Nee	NTB	Onbeslist	Onbeslist	NTB	Nee	Nee	Nee	Nee	Geen data
		Negatief	Positief	Negatief	Negatief	Nee	Nee	NTB	Nee	Nee	Nee	Onbeslist	NTB	Nee	Negatief
		Nee	Nee	Nee	NTB	Nee	Negatief	Negatief	Nee	NTB	Nee	Nee	Nee	Nee	Negatief
		Nee	Nee	Onbeslist	Onbeslist	Nee	NTB	Onbeslist	Onbeslist	NTB	Nee	Negatief	Nee	Onbeslist	Onbeslist
		Nee	Nee	NTB	Nee	Nee	Onbeslist	Nee	Nee	Nee	Nee	Negatief	NTB	Nee	Onbeslist
		Nee	Nee	Nee	NTB	Positief	Negatief	Negatief	Positief	NTB	Negatief	Negatief	Nee	Positief	Nee
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Onbeslist	Nee	Nee	Nee	Onbeslist	NTB	Nee	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Onbeslist	Nee	Nee	Nee	Onbeslist	NTB	Nee	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Negatief	Negatief	Nee	NTB	Negatief	Nee	Nee	Positief	Negatief
		Negatief	Positief	Nee	NTB	Positief	Negatief	Negatief	Onbeslist	NTB	Nee	Negatief	Nee	Onbeslist	Onbeslist
		Nee	Nee	Onbeslist	Onbeslist	Nee	NTB	Onbeslist	Onbeslist	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Onbeslist	Nee	Onbeslist	Nee	Onbeslist	NTB	Nee	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Onbeslist	Nee	Negatief	Nee	Onbeslist	NTB	Nee	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Negatief	Negatief	Nee	NTB	Negatief	Negatief	Nee	Onbeslist	Nee
		Negatief	Positief	Negatief	NTB	Nee	Negatief	Negatief	Nee	NTB	Nee	Negatief	Negatief	Nee	Onbeslist
		Negatief	Positief	NTB	Positief	Positief	Negatief	Positief	Positief	NTB	Nee	Negatief	Nee	Positief	Onbeslist
		Nee	Nee	NTB	Nee	Onbeslist	Onbeslist	Nee	Positief	Nee	NTB	Onbeslist	Nee	Onbeslist	Onbeslist
		Nee	Nee	Nee	NTB	Positief	Negatief	Nee	Nee	Negatief	Negatief	Negatief	Nee	Nee	Nee
		Nee	Nee	Nee	NTB	Positief	Negatief	Onbeslist	Negatief	NTB	Negatief	Negatief	Nee	Negatief	Onbeslist
		Nee	Nee	Nee	Nee	Nee	NTB	Onbeslist	Nee	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Nee	Geen data	Geen data	Nee	NTB	Onbeslist	Onbeslist	NTB	Nee	Onbeslist	Nee	Onbeslist	Geen data
		Nee	Nee	Negatief	NTB	Nee	Nee	Nee	Onbeslist	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Onbeslist	NTB	Positief	Positief	Onbeslist	Onbeslist	Onbeslist	NTB	Positief	Onbeslist	Positief	Onbeslist	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Nee	Onbeslist	Onbeslist	NTB	Positief	Positief	Positief	Onbeslist	Onbeslist
		Nee	Nee	Nee	NTB	Nee	Negatief	Nee	Nee	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Nee	Geen data	Geen data	Negatief	NTB	Onbeslist	Onbeslist	NTB	Nee	Nee	Nee	Onbeslist	Onbeslist
		Nee	Nee	Negatief	NTB	Nee	Negatief	Onbeslist	Negatief	NTB	Negatief	Negatief	Nee	Negatief	Onbeslist
		Nee	Nee	Negatief	NTB	Nee	Negatief	Onbeslist	Onbeslist	NTB	Negatief	Nee	Negatief	Onbeslist	Onbeslist
		Negatief	Positief	Nee	NTB	Positief	Negatief	Negatief	Onbeslist	NTB	Nee	Nee	Nee	Onbeslist	Nee
		Negatief	Positief	Negatief	NTB	Positief	Nee	Positief	Negatief	NTB	Negatief	Negatief	Nee	Negatief	Onbeslist
		Negatief	Positief	Negatief	NTB	Positief	Nee	Positief	Positief	Nee	Negatief	Negatief	NTB	Onbeslist	Onbeslist
		Nee	Nee	NTB	Nee	Positief	Onbeslist	Nee	Onbeslist	Nee	Nee	Onbeslist	NTB	Onbeslist	Onbeslist
		Nee	Nee	Nee	NTB	Nee	Nee	Negatief	Nee	NTB	Nee	Negatief	Nee	Nee	Negatief

Figuur 4.1: Fairness conclusies per variabele voor alle Ad en Bachelor opleidingen van De Haagse Hogeschool

4.2 TOEPASSINGEN VAN DIT ONDERZOEK

Nu ik de belangrijkste inzichten heb besproken, ben ik toegekomen aan het tweede aspect van toegepast onderzoek: wat hebben we eraan in de praktijk? Er zijn meerdere onderzoeksproducten uit dit onderzoek die direct nuttig zijn:

- **Een instrument om kansengelijkheid te onderzoeken binnen De HHs en de regio** – Met dit onderzoek heeft De Haagse een instrument in handen om te bepalen of er groepen studenten zijn die onevenredig minder kansen hebben in hun studie, en daar vervolgens mee aan de slag te gaan. Het uitgevoerde onderzoek biedt de mogelijkheid om dit te onderzoeken binnen de regio. Dit sluit perfect aan bij de agenda van de Gelijke Kansen Alliantie van de gemeente Den Haag (2019-2021) (*GKA-Agenda gemeente Den Haag, 2019*).

“Alle kinderen en jongeren hebben recht op gelijke kansen in het onderwijs. Het is belangrijk dat de talenten van alle kinderen en jongeren optimaal benut worden. Dat is niet alleen van groot belang voor de toekomst van kinderen en jongeren zelf, maar ook voor de toekomst van onze samenleving.”

De eerste gebieden die we regionaal willen toetsen zijn de doorstroom van mbo naar hbo en stagediscriminatie.

- **Een landelijke, statistische onderzoeksmethode naar kansengelijkheid in het vervolgonderwijs** – Daarnaast biedt dit onderzoek een instrument aan alle onderwijsinstellingen om kansengelijkheid kwantitatief te onderzoeken. Ik ben blij te kunnen melden dat we deze methode als open access en open source ter beschikking zullen stellen. We hebben Npuls bereid gevonden om de opschaling van de methode en achterliggende broncode te adopteren via het Center for Educational Data Analytics. Vanaf 1 december zal ik hier tot de zomer van 2025 aan gaan werken met de collega's van Npuls. Meer informatie hierover komt beschikbaar via de websites van Npuls en De Haagse Hogeschool.
- **Een synthetische studiedataset waarmee studenten onderzoek kunnen doen** – Tot slot hebben we voor de uitleg van de methode een synthetische dataset ontwikkeld op basis van bestaande opleidingen van De Haagse Hogeschool. Deze dataset heeft de kenmerken van alle opleidingen met inbegrip van de verdeling van variabelen, maar is wel volledig fictief. Studenten van De Haagse kunnen deze set gebruiken om zelf onderzoek te doen met het lectoraat.

4.3 MOGELIJKE TEKORTKOMINGEN, RISICO'S EN AANBEVELINGEN

Onderzoek naar algoritmische fairness in het hoger onderwijs leert dat dit type onderzoek zich ontwikkelt van onbekende naar bekende bias (Baker & Hawn, 2022). Ons onderzoek draagt hier verder aan bij. Maar net als elk onderzoek kent ook ons

onderzoek ook tekortkomingen en risico's. Ik bespreek deze en geef ook aanbevelingen om die tegen te gaan.

- **De aanpak houdt geen rekening met individuele verschillen** – In de opbouw van fairness analyse onderzoeken we wel de mate van variantie binnen groepen, maar houden we geen rekening met individuele fairness. Het is mogelijk dat de uitspraken die we doen over groepen, niet opgaan voor individuen binnen die groepen, doordat er andere, onbekende kenmerken zijn die in hun voor- of nadeel werken.

Mijn *aanbeveling* is de methode te verfijnen met overstijgende analyses, bijvoorbeeld op combinaties van vooropleiding, geslacht en soort instroom. Houd verder altijd rekening met de persoonlijke situatie van de student.

- **De aanpak houdt geen rekening met bijzondere persoonsgegevens** – Toevoeging van bijzondere persoonsgegevens, zoals data over migratieachtergrond of ondersteuningsbehoefte, kan aan fairness analyses nog meer context geven.

Mijn *aanbeveling* is deze gegevens te betrekken in dit type onderzoek. In het huidige onderzoek heb ik deze gegevens bewust buiten scope geplaatst. Niet omdat ik ze niet van belang vind – juist wel –, maar omdat het voor onze hogeschool nieuw is om op grote schaal data te verzamelen en te combineren voor dit type onderzoek. Daarnaast speelt mee dat we vanwege het *informed consent* dat voorwaardelijk is voor de verwerking van bijzondere persoonsgegevens een nieuw type bias introduceren, namelijk *consent bias*. Met name studenten die meer vertrouwen hebben in een onderwijsinstelling zijn ook meer geneigd om toestemming te geven voor het verzamelen van dit type data (Li et al., 2022).

- **Het onderzoek geeft nog geen handvatten om een eventueel gebrek aan kansengelijkheid op te lossen** – Deze methode legt wel de vinger op de zere plek, maar geeft niet aan waar eventueel een oplossing zit. Voor alle duidelijkheid – dat is niet de aanpassing van het voorspelmodel of de studiedata. Omdat deze methode inzicht geeft in fairness per groep, moeten oplossingen gezocht worden in mogelijke interventies die voor specifieke groepen studenten baat kunnen hebben.

Docenten, begeleiders en bestuurders zijn gebonden aan een *decision space*. Alternatieve opties zijn uitgesloten of slechts met moeite bereikbaar, bijvoorbeeld uitstel of afschaffing van bsa, alternatieve toetsing, vrijstellingen, extra tijd, etc. Ten tijde van corona waren deze opties wel beschikbaar, wat het interessant

maakt juist die te bestuderen. Onderzoek naar verschillen in studiesucces tussen studenten voor en na corona (Vooren et al., 2024) laat zien dat afschaffing van het bindend studieadvies al zou bijdragen aan kansengelijkheid in retentie en diplomasucces. Hoe zou die bredere decision space eruit zien als andere maatregelen ook werden toegepast op alle studenten, zoals extra tijd voor tentamens, een lesverlaatkaart, verschillende toetsvormen of een alternatieve invulling voor bepaalde eindtermen?

Mijn *aanbeveling* is om verder in overleg met docenten en studenten vast te stellen of er sprake is van kansengelijkheid die de opleiding kan en wil aanpakken. Ik pleit daarbij voor groepsarrangementen en niet voor individuele toepassing van voorspelmodellen.

Allereerst moet bepaald worden in welke mate een oplossing gevonden kan worden in het onderwijs. Dit kan door individuele vakken in het curriculum nader te onderzoeken. Misschien blijkt hieruit dat bepaalde vaardigheden in onvoldoende mate aanwezig zijn. Hierop kan bijvoorbeeld gericht deficiëntie-onderwijs aangeboden worden voor de gehele studentenpopulatie, zodat studenten hier vrijwillig aan kunnen deelnemen. Mocht blijken uit bijvoorbeeld panel-interviews dat eventuele barrières zich bevinden in de meer persoonlijke levenssfeer, kan daarop in de zorgstructuur misschien naar oplossingen worden gezocht. Een verschil met de huidige praktijk is, dat oplossingen niet pas gezocht worden nadat er problemen zijn die via examencommissies, begeleiders of psychologen naar voren komen, maar dat deze proactief worden aangeboden.

- **Een gebrek aan gelijke kansen staat niet gelijk aan discriminatie** – Er bestaat het risico dat we de uitkomst over kansengelijkheid verwarren met discriminatie. Dit onderzoek had tot doel om gelijke kansen te onderzoeken in de context van het hoger onderwijs en De Haagse Hogeschool in het bijzonder; dit kan los staan van discriminatie. Discriminatie is “mensen anders behandelen, achterstellen of uitsluiten op basis van persoonskenmerken” (“Wat is discriminatie? | Mensenrechten voor jou | College voor de Rechten van de Mens,” n.d.). Het is de vraag of eventuele achterstelling *moedwillig* wordt gemaakt. Dit zal verwacht ik niet het geval zijn.

Mijn *aanbeveling* is om te onderzoeken of studenten de werking van het huidige onderwijssysteem en bijbehorende procedures als discriminerend *ervaren*. Binnen het Centre of Expertise Global & Inclusive Learning zijn er meerdere collega’s die in dit type onderzoek zijn gespecialiseerd.

- **Deze inzichten kunnen gebruikt worden voor het wegnemen van bottlenecks, maar ook voor het verlagen daarvan en het opwerpen van extra barrières** – De neiging zou kunnen ontstaan om inzichten uit dit onderzoek te gebruiken om studenten die in het huidige onderwijssysteem minder kans hebben op succes te ontmoedigen. Dit zou ogenschijnlijk kunnen bijdragen aan wat sommigen zien als de kwaliteit van de opleiding. Anders dan algoritmes hebben mensen *wel* een ethisch kompas waarop ik hier een sterk beroep doe.

Mijn *aanbeveling* is dat dit appèl wordt voorzien van een gedragscode Privacy, ethiek en AI voor onze hogeschool. Dit sluit goed aan bij de AI Act die per 1 augustus van dit jaar van kracht is geworden die het onderwijs kenmerkt als een hoog risico gebied waarvoor extra waarborgen gelden (Engelfriet, 2024). Een daarvan is dat onderwijsinstellingen aantonen of eventuele algoritmes rekening houden met bias en dat zij medewerkers trainen op het herkennen en tegengaan van bias.

- **Deze voorspelmodellen mogen niet gebruikt worden voor individuele toepassing op huidige of aankomende studenten** – In lijn met het vorige risico bestaat er misschien de wens om deze voorspelmodellen te gebruiken om individueel toe te passen. Zelfs als dit met *informed consent* zou zijn van de student, is dit niet toegestaan. Ze zijn uitsluitend ontwikkeld om de mate van fairness vast te stellen. Op dit punt is mijn *aanbeveling* dat voor de ontwikkeling van voorspelmodellen binnen het onderwijs of de begeleiding de samenwerking met het lectoraat wordt opgezocht om hierin rekening te houden met eventuele bias en – samen met studenten – nieuwe instrumenten verder te ontwikkelen.

4.4 VERVOLGONDERZOEK

Ik rond graag af met het vervolgonderzoek dat we willen uitvoeren door verbreding en verdieping.

- **Andere succesvariabelen** — Naar andere succesvariabelen, zoals de kans op het behalen van een diploma. Het is vrijwel zeker dat de langstudeermaatregel die het nieuwe kabinet heeft afgekondigd nadelig zal zijn van kansengelijkheid in het onderwijs. Juist voor studenten die meer tijd nodig hebben om het hoger onderwijs te bereiken, is het waarschijnlijk dat zij ook meer tijd nodig hebben om dat af te ronden.

- **Bijzondere persoonsgegevens** – Verder willen we onderzoeken wat de invloed is van het betrekken van bijzondere persoonsgegevens met informed consent.
- **Onderdelen van curricula** – Daarnaast zullen we in de doorstroom ook kansengelijkheid onderzoeken binnen het curriculum van opleidingen: individuele vakken, groepen van vakken en grotere studieonderdelen. Denk hierbij in het bijzonder aan mogelijke stagediscriminatie.
- **Arbeidsmarktsucces** — Verder zijn we in gesprek met de Randstad 6 hogescholen om te bezien of we met elkaar een kansengelijkheidsonderzoek kunnen doen onder alle studenten die met of zonder een afgeronde opleiding aan onze hogescholen hebben gestudeerd. Daarvoor willen we onderzoek doen op basis van CBS microdata.
- **Replicatie** - Daarnaast hopen we dat ons onderzoek gerepliceerd kan worden aan andere onderwijsinstellingen. De opschaling vanuit het Npuls programma zal hier zeker bij gaan helpen.
- **Learning Technology** – Tot slot zijn we binnen *learning technology* twee onderzoekslijnen gestart naar effecten van interventies, in het bijzonder bij de inzet van EdTech. En daarnaast naar *honesty* in online toetsen. Met de komst van een nieuwe senior onderzoeker op dit onderwerp verwachten we hier een prachtige invulling aan te kunnen geven.

Hiermee ben ik aan het eind gekomen van mijn intreerede. Ik heb u een impressie gegeven van mijn lectoraat en de potentie van dit type onderzoek. Ik hoop dat ik hiermee een wezenlijke bijdrage kan leveren aan het vakgebied van learning analytics in Nederland en aan kansengelijkheid en sociale rechtvaardigheid in het Nederlands onderwijsstelsel voor al onze studenten.

Let's Change. You. Us. The World.

SAMENVATTING

*Sometimes it is the people no one imagines anything of
who do the things that no one can imagine.*

— A. Turing

In zijn intrede *No Fairness without Awareness – Toegepast onderzoek naar kansengelijkheid in het hoger onderwijs* introduceert Dr. Theo C. Bakker het thema kansengelijkheid in het hoger onderwijs en de rol van learning analytics om sociale rechtvaardigheid te bevorderen. Vanuit zijn lectoraat Learning Technology & Analytics aan De Haagse Hogeschool heeft Bakker een statistisch onderzoek uitgevoerd naar kansengelijkheid in het hoger beroepsonderwijs (hbo), waarin hij met behulp van studiedata en AI-voorspelmodellen onderzoekt in hoeverre verschillende studentengroepen gelijk behandeld worden. Hij benadrukt dat kansengelijkheid niet alleen gaat om toegang tot het onderwijs, maar ook om eerlijke kansen op studiesucces en doorstroom naar de arbeidsmarkt.

BELANG VAN LEARNING ANALYTICS VOOR KANSENGELIJKHEID

Learning analytics omvat het meten, verzamelen en analyseren van data over studenten en hun leeromgeving, met als doel het onderwijs en de ondersteuning van studenten te optimaliseren. In het onderzoek maakt Bakker daarbinnen een nader onderscheid tussen verschillende soorten analyses: *learning analytics*, dat de interacties tussen studenten, docenten en de lesstof onderzoekt; *student analytics*, waarbij individuele studentgegevens worden onderzocht; *institutional analytics*, dat kijkt naar gegevens op organisatieniveau; en *inclusion analytics*, dat zich richt op sociale rechtvaardigheid en kansengelijkheid. Bij het laatste type onderzoekt hij of bepaalde groepen studenten stelselmatig benadeeld worden, dat bijvoorbeeld blijkt uit een grotere kans op uitval.

Bakker bespreekt daarvoor het probleem van bias in onderwijsdata. Bias is een systematische fout in de data die kan leiden tot oneerlijke uitkomsten. Algoritmen versterken vaak bestaande vooroordelen in de data, wat in het onderwijs kan leiden tot kansengelijkheid. Hij benadrukt dat kansengelijkheid alleen kan worden gerealiseerd

als we ons bewust zijn van mogelijke onrechtvaardigheden in het systeem: *No Fairness without Awareness*.

METHODE: ANALYSE VAN BIAS EN FAIRNESS

Voor zijn onderzoek maakt Bakker gebruik van machine learning-algoritmen die retentie na het eerste studiejaar voorspellen. Door deze voorspellingen te analyseren, onderzoekt hij of de kansen eerlijk zijn verdeeld tussen groepen, bijvoorbeeld tussen mannen en vrouwen of tussen mbo-, havo- en vwo-studenten. Een belangrijke maatstaf in het onderzoek is het 4/5-criterium, waarbij wordt beoordeeld of de kansen voor de 'beschermdde groep' ten minste 80% bedragen van die van de 'bevoorrechte groep'. Als de verhouding kleiner is, is er sprake van een gebrek aan fairness.

Bakker beschrijft drie stappen in de analyse:

1. *Data voorbereiden en voorspelmodellen ontwikkelen*: De data worden opgeschoond en opgesplitst in een trainings-, test-, en validatieset. Modellen zoals logistische regressie en random forest worden gebruikt om te voorspellen welke studenten waarschijnlijk zullen doorstromen.
2. *Bepalen van het belang van variabelen*: Er wordt gekeken naar de invloed van kenmerken zoals vooropleiding en leeftijd op de uitkomsten. Voor verschillende groepen wordt een fictieve 'gemiddelde student' gemodelleerd om inzicht te krijgen in de verschillen in kansen.
3. *Fairness berekenen*: Met het 4/5-criterium worden verschillen tussen groepen vergeleken. Dit criterium wordt toegepast op verschillende *metrics* uit voorspelmodellen zoals accuraatheid, gelijke kansen en statistische gelijkheid. De analyse toont aan waar bias optreedt en hoe deze de kansengelijkheid beïnvloedt.

RESULTATEN EN CONCLUSIES

Het onderzoek van Bakker toont aan dat op veel opleidingen in het hbo sprake is van ongelijkheid in kansen. Zo hebben bij De Haagse Hogeschool vrouwen in slechts één opleiding een negatieve bias, terwijl mannen in twintig opleidingen een lagere kans op retentie hebben. Vwo-studenten hebben vaker een positieve bias, terwijl mbo-studenten juist benadeeld worden. De aanpak biedt concrete inzichten voor onderwijsinstellingen om ongelijkheden te identificeren en te analyseren. Dit kan bijdragen aan interventies

die specifiek gericht zijn op het ondersteunen van groepen studenten die onevenredig meer risico lopen op uitval. De methodiek kan bovendien opgeschaald worden en biedt een herhaalbare aanpak om bias in studiedata te detecteren. Deze opschaling zal binnen het landelijke Npuls groeifondsprogramma worden opgepakt.

AANBEVELINGEN EN VERVOLGONDERZOEK

Bakker doet verschillende aanbevelingen om kansengelijkheid in het hoger onderwijs te bevorderen. Ten eerste pleit hij voor het gebruik van overstijgende analyses om complexere patronen van ongelijkheid beter te begrijpen. Daarnaast raadt hij aan om sensitieve kenmerken, zoals migratieachtergrond en sociaaleconomische status, op een verantwoorde manier mee te nemen in analyses, zodat een volledig beeld van de ongelijkheden ontstaat. Hij waarschuwt wel voor *consent bias*: studenten die vertrouwen hebben in de onderwijsinstelling geven eerder toestemming voor het gebruik van persoonlijke data, wat de representativiteit van de data kan beïnvloeden. Een andere aanbeveling betreft het ontwikkelen van interventies om knelpunten te verminderen. Bakker suggereert dat opleidingsmanagers en docenten in onderwijsinstellingen een grotere *decision space* moeten krijgen, met meer mogelijkheden voor flexibele en proactieve ondersteuning van studenten. Hij verwijst naar maatregelen tijdens de coronacrisis, zoals het uitstellen van het bindend studieadvies, die hebben aangetoond dat flexibiliteit in beleid de kansen op studiesucces vergroot. Ten slotte stelt Bakker dat discriminatie en kansengelijkheid verschillende fenomenen zijn. Niet elke ongelijkheid in uitkomsten is een gevolg van discriminatie, maar kan ook voortkomen uit systeemfouten in het onderwijs zelf.

TOEKOMSTPERSPECTIEF

Het onderzoek van Bakker levert waardevolle inzichten op voor beleidsmakers, docenten en bestuurders in het hoger onderwijs. Zijn aanbevelingen geven richting aan nieuwe studies en beleidsmaatregelen die kansengelijkheid kunnen verminderen. In vervolgonderzoek wil hij niet alleen naar retentie, maar ook naar het behalen van diploma's en stagediscriminatie kijken. Het lectoraat zal de gebruikte methodiek via Npuls open access beschikbaar stellen, waardoor andere instellingen de aanpak kunnen overnemen en verbreden.

Bakker besluit zijn intreerede met de oproep om data-geïnformeerde beslissingen te gebruiken voor een sociaal rechtvaardiger onderwijssysteem, dat studenten een eerlijke kans biedt om te slagen.

Let's Change. You. Us. The World.

SUMMARY

*Sometimes it is the people no one imagines anything of
who do the things that no one can imagine.*

— A. Turing

In his inaugural address, *No Fairness without Awareness – Applied Research on Equity in Higher Education*, Dr Theo C. Bakker introduces the topic of equity in higher education and the role of learning analytics in promoting social justice. From his professorship in Learning Technology & Analytics at The Hague University of Applied Sciences (THUAS), Bakker has been conducting statistical research on equity in higher vocational education (hbo), using student data and AI predictive models to examine the extent to which different student groups are treated equally. He emphasizes that equity of opportunity is about access to education and fair opportunities for study success and advancement into the labor market.

IMPORTANCE OF LEARNING ANALYTICS FOR EQUITY

Learning analytics involves measuring, collecting, and analyzing data about students and their learning environment to optimize instruction and support for students. Within this field, Bakker further distinguishes between different types of analytics: *learning analytics*, which examines the interactions between students, teachers and the subject matter; *student analytics*, which examines individual student data; *institutional analytics*, which looks at organizational-level data; and *inclusion analytics*, which focuses on social justice and equity. In the latter type, he examines whether certain groups of students are systematically disadvantaged, for instance to drop out.

To that end, Bakker discusses the problem of bias in educational data. Bias is a systematic error in data that can lead to unfair outcomes. Algorithms often reinforce existing biases in the data, leading to inequity in education. He emphasizes that equity can only be realized if we are aware of possible injustices in the system: *No Fairness without Awareness*.

METHOD: ANALYSIS OF BIAS AND FAIRNESS

For his research, Bakker uses machine learning algorithms that predict retention after the first year of study. By analyzing these predictions, he examines whether the odds are fairly distributed between groups, for example, between men and women or between mbo-, havo- en vwo-students. An essential measure in the study is the 4/5 criterion, which assesses whether the odds for the 'protected group' are at least 80% of those of the 'advantaged group'. If the ratio is smaller, there is a lack of fairness.

Bakker describes three steps in the analysis:

1. *Prepare data and develop prediction models*: The data are cleaned and split into a training, testing, and validation set. Models such as logistic regression and random forest are used to predict which students are likely to pass the first year.
2. *Determining the importance of variables*: The influence of characteristics such as prior education and age on outcomes is examined. A fictitious 'average student' is modelled for different groups to understand differences in odds.
3. *Calculating Fairness*: The 4/5 criterion is used to compare differences in fairness between groups. This criterion is applied to various metrics from prediction models, such as accuracy, equality of opportunity and statistical equality. The analysis shows where bias occurs and how it affects equity.

RESULTS AND CONCLUSIONS

Bakker's research shows that inequity exists at the majority of programs. At THUAS, for example, women in only one program have a negative bias, while men in 20 programs have a lower chance of retention. Vwo-students are more likely to have a positive bias, while mbo-students are disadvantaged. The approach provides concrete insights for educational institutions to identify and analyze inequities. This approach can contribute to interventions specifically aimed at supporting groups of students with a disproportional higher risk of dropping out. Moreover, the methodology can be scaled up and provides a replicable approach to detect bias in student data. A national scale-up will be funded by the Npuls growth fund program.

RECOMMENDATIONS AND FOLLOW-UP RESEARCH

Bakker makes several recommendations to promote equity in higher education. First, he advocates using cross-cutting analyses to improve understanding of more complex inequality patterns. In addition, he recommends responsibly including sensitive characteristics, such as migration background and socioeconomic status, in analyses to provide a fuller picture of inequalities. He warns of *consent bias*: students who trust the educational institution are more likely to consent to the use of personal data, which can affect the representativeness of the data.

Another recommendation concerns developing interventions to reduce bottlenecks. Bakker suggests that program managers and teachers in educational institutions should be given more significant *decision space*, with more opportunities for flexible and proactive support for students. He refers to measures during the Covid crisis, such as the postponement of the binding study advice, which have shown that flexibility in policy increases the chances of study success. Finally, Bakker argues that discrimination and inequity are different phenomena. Not every inequality in outcomes results from discrimination; it can also stem from systemic failures in education.

FUTURE PROSPECT

Bakker's research provides valuable insights for policymakers, educators and administrators in higher education. His recommendations guide new studies and policies that can reduce inequity. In follow-up research, he plans to look at retention, degree attainment, and internship discrimination. The professorship will make the methodology available through Npuls open access, allowing other institutions to adopt and broaden the approach.

Bakker concludes his inaugural address by calling for data-informed decisions to be used for a more socially just educational system that gives students a fair chance to succeed.

Let's Change. You. Us. The World.

DANKWOORD

*Learn from yesterday, live for today, hope for tomorrow.
The important thing is not to stop questioning.*

— A. Einstein

Ik wil het College van Bestuur van De Haagse Hogeschool, de directie van de faculteit IT & Design en de lectoren van het Centre of Expertise Global & Inclusive Learning van harte bedanken voor de kans die zij mij hebben geboden om lector Learning Technology & Analytics te worden. Ik hoop dat ik de belofte heb kunnen waarmaken. Dank aan mijn oud collega's van de dienst Onderwijs, Kennis & Communicatie (OKC) die na mijn vertrek als directeur de taken hebben waargenomen en overgenomen.

Binnen de hogeschool wil ik verder alle collega's bedanken die dit onderzoek mogelijk hebben gemaakt: de collega's van Juridische Zaken voor het meewerken aan de DPIA voor dit onderzoek, de collega's van de hogeschoolbibliotheek voor de support bij het datamanagement, de leden van de ethische commissie voor het bevragen van het onderzoeksvoorstel en de collega's van OCK, team IR & Analytics, voor het leveren van de onderzoeksdata. Buiten de hogeschool wil ik de collega's van Npuls bedanken dat zij mij de kans geven om dit onderzoek nationaal verder op te schalen.

Ook veel dank aan alle leden van de kenniskring Learning Technology & Analytics voor hun enthousiasme voor ons onderzoek naar Learning Analytics of Learning Technology. Ik leer elke dag van jullie. Daarnaast veel dank aan mijn geweldige managementassistente, Titia Dijkstra. Zonder jou zou het lectoraat nergens zijn! Dear Manika, Thank you for joining my research group. It is a delight working with you.

Tot slot veel dank voor alle steun die ik thuis heb gekregen van mijn lieve echtgenoot om deze stap te maken. Lieve Sjors, waar ik ondernemer was toen we elkaar leerden kennen en jij onderzoeker, hebben we nu van plek gewisseld. Ik wil mijn ouders bedanken voor de liefde voor onderwijs en studeren, die me met de paplepel is ingegoten.

REFERENTIES

- Badou, M., & Day, M. (2021, October). *Kansengelijkheid in het onderwijs. Verkennend onderzoek naar factoren die samenhangen met onderwijs(on)gelijkheid* (tech. rep.). Gelijke kansen alliantie. Retrieved October 28, 2022, from <https://www.gelijkekansen.nl/documenten/publicaties/2021/10/05/verkennd-onderzoek-naar-factoren-die-samenhangen-met-kansengelijkheid>
- Baker, R. S., & Hawn, A. (2022). Algorithmic Bias in Education. *International Journal of Artificial Intelligence in Education*, 32(4), 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Barocas, S., Hardt, M., & Narayanan, A. (2023, December). *Fairness and Machine Learning: Limitations and Opportunities*. fairmlbook.org. Retrieved June 22, 2024, from <http://www.fairmlbook.org>
- Baum, S., Ma, J., & Payea, K. (2010). *Education Pays 2010: The Benefits of Higher Education for Individuals and Society* (tech. rep.). CollegeBoard. <https://files.eric.ed.gov/fulltext/ED526357.pdf>
- Belfi, B., Allen, J., Jacobs, B., & Oosterhout, K. V. (2018). *Arbeidsmarktongelijkheid onder hbo'ers: Hoe individuele achtergrondkenmerken van hbo'ers samenhangen met arbeidsmarktsucces op de korte en middellange termijn* (tech. rep.). Retrieved November 21, 2023, from https://www.hbomonitor.nl/application/files/1915/4503/3607/ROA%5C_R%5C_2018%5C_9.pdf
- Biecek, P. (2018). Dalex: Explainers for complex predictive models in r. *Journal of Machine Learning Research*, 19(84), 1–5. <https://jmlr.org/papers/v19/18-416.html>
- Boer, T. d. (2024). Evidence-informed master admission.
- Code of Federal Regulations. Section 4d, uniform guidelines on employee selection procedures (1978). (1978). Retrieved May 19, 2023, from <https://www.govinfo.gov/content/pkg/CFR-2011-title29-vol4/xml/CFR-2011-title29-vol4-part1607.xml>
- Copier, J. (2022). *Tussen idealen en dwalingen. Verhalen over onderwijs*. Garant.
- Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2011). Fairness Through Awareness. *arXiv*. <https://doi.org/10.48550/arxiv.1104.3913>

- Elffers, L. (2022, May). *Onderwijs maakt het verschil - kansengelijkheid in het Nederlandse onderwijs*. Walburg Pers B.V.
- Engelfriet, A. (2024, January). *AI and algorithms Mastering legal and ethical compliance*. Ius Mentis (ICTRecht).
- Ersoy, S., & Gaag, S. V. d. (2023, June). Studenten met migratieachtergrond opvallend vaak beschuldigd van fraude, minister wil systeem grondig nagaan. Retrieved June 21, 2023, from <https://nos.nl/op3/artikel/2479700-studenten-met-migratieachtergrond-opvallend-vaak-beschuldigd-van-fraude-minister-wil-systeem-grondig-nagaan>
- Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4615421>
- GKA-Agenda gemeente Den Haag (tech. rep.). (2019, July). Gemeente Den Haag. Retrieved August 1, 2024, from <https://www.gelijke-kansen.nl/gemeenten/den-haag>
- Handboek over het Europese non-discriminatierecht, Editie 2018*. (2018). https://fra.europa.eu/sites/default/files/fra%5C_uploads/fra-2018-handbook-non-discrimination-law-2018%5C_nl.pdf
- Hardt, M., Price, E., & Srebro, N. (2016). Equality of Opportunity in Supervised Learning. *arXiv*. <https://doi.org/10.48550/arxiv.1610.02413>
- Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles*. <https://www.tidymodels.org>
- Li, W., Sun, K., Schaub, F., & Brooks, C. (2022). Disparities in Students' Propensity to Consent to Learning Analytics. *International Journal of Artificial Intelligence in Education*, 32(3), 564–608. <https://doi.org/10.1007/s40593-021-00254-2>
- Ma, J., Pender, M., & Welch, M. (2016). *Education Pays 2016: The Benefits of Higher Education for Individuals and Society*. (tech. rep.). CollegeBoard. <https://files.eric.ed.gov/fulltext/ED572548.pdf>
- Maksymiuk, S., Gosiewska, A., & Biecek, P. (2020). Landscape of r packages for explainable artificial intelligence [Pages 6, 7, 11, 15]. *arXiv*. <https://arxiv.org/abs/2009.13248>
- Murray, J. (2009). The Wider Social Benefits of Higher Education: What do We Know about Them? *Australian Journal of Education*, 53(3), 230–244. <https://doi.org/10.1177/000494410905300303>
- Onderzoekend leren met impact. (2022). <https://www.dehaagsehogeschool.nl/sites/hhs/files/documents/OverDeHaagse-Organisatie-FeitenDocumenten-instellingsplan-02-15-def.pdf>

- O'Neil, C. (2017, April). The era of blind faith in big data must end. Retrieved July 29, 2024, from [The%20era%20of%20blind%20faith%20in%20big%20data%20must%20end](#)
- Pedreshi, D., Ruggieri, S., & Turini, F. (2008). Discrimination-aware data mining. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. <https://doi.org/10.1145/1401890.1401959>
- Perna, L. W. (2005). The Benefits of Higher Education: Sex, Racial/Ethnic, and Socioeconomic Group Differences. *The Review of Higher Education*, 29(1), 23–52. <https://doi.org/10.1353/rhe.2005.0073>
- Position paper Gelijke kansen voor de hogeschoolstudent* (tech. rep.). (2023, February). Vereniging Hogescholen.
- Professionals voor morgen - Strategische agenda, Vereniging Hogescholen 2019-2023* (tech. rep.). (2019, September). Vereniging Hogescholen. Den Haag. Retrieved July 27, 2024, from https://www.vereniginghogescholen.nl/system/knowledge%5C_base/attachments/files/000/001/067/original/Professionals%5C_voor%5C_morgen%5C_-%5C_strategische%5C_agenda%5C_Vereniging%5C_Hogescholen%5C_2019-2023.pdf
- Rutgers, M., Kuijk, M. v., Raad, P. v., & Ovbiagbonhia, R. (2022, October). *Onderzoek studentbehoefte flexibel onderwijs. De eerste Hanze-verkenning* (tech. rep.). Retrieved August 1, 2024, from https://research.hanze.nl/ws/portalfiles/portal/58916269/1.%5C_Rapportage%5C_onderzoek%5C_studentbehoefte%5C_flexibel%5C_onderwijs%5C_2022.pdf
- SDG Nederland - 4 Kwaliteitsonderwijs. (n.d.). Retrieved July 27, 2024, from <https://www.sdg nederland.nl/SDG/4-kwaliteitsonderwijs/>
- Sectorbeeld Onderwijs 2021-2023* (tech. rep.). (2024, January). Autoriteit Persoonsgegevens. Retrieved August 1, 2024, from <https://www.autoriteitpersoonsgegevens.nl/documenten/sectorbeeld-onderwijs-2021-2023>
- Shin, D. (2022). How do people judge the credibility of algorithmic sources? *AI & SOCIETY*, 37(1), 81–96. <https://doi.org/10.1007/s00146-021-01158-4>
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-sts330>
- Shmueli, G., & Koppius, O. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, 35(3), 553. <https://doi.org/10.2307/23042796>
- Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, (46), 31–40.

- Smeets, E., Geurts, R., & Helvoirt, D. v. (2024). *Algoritmen in het onderwijs. Een onderzoek in opdracht van het College voor de Rechten van de Mens* (tech. rep.). KBA Nijmegen / ResearchNed.
- Stift, R. (2024, November). Tienduizend studenten krijgen compensatie wegens discriminatie DUO. Retrieved November 11, 2024, from <https://www.nrc.nl/nieuws/2024/11/11/tienduizend-studenten-krijgen-compensatie-wegens-discriminatie-duo-a4872677>
- Taakopdracht sectorale verkenning hoger economisch onderwijs (HEO) 2020-2021* (tech. rep.). (2020). Vereniging Hogescholen. Retrieved July 30, 2024, from https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/001/127/original/Definitieve_taaopdracht_sectorale_verkenning_heo_2020_-_2021.pdf?1594373979
- Talent op de juiste plek* (tech. rep.). (2024, July). IBO doelmatig Hoger onderwijs. Retrieved July 26, 2024, from https://www.tweedekamer.nl/kamerstukken/brieven%5C_regering/detail?did=2024D29131%5C&id=2024Z12068
- Tinto, V. (2012, March). *Completing College: Rethinking Institutional Action*. University of Chicago Press.
- Vervolgonderzoek bevestigt indirecte discriminatie controles uitwonendenbeurs. (2024, May). Retrieved August 1, 2024, from <https://www.duo.nl/particulier/home/actueel/vervolgonderzoek-bevestigt-indirecte-discriminatie-controles-uitwonendenbeurs.jsp>
- Vooren, M., Cornelisz, I., Meeter, M., & Klaveren, C. V. (2024). Bindend studieadvies helpt studenten niet. *Economisch Statistische Berichten*, 4836(109), 351–353.
- Vrijwel geen toezicht op gebruik algoritmes in het onderwijs. (2024, May). Retrieved August 1, 2024, from <https://www.scienceguide.nl/2024/05/vrijwel-geen-toezicht-op-gebruik-algoritmes-in-het-onderwijs/>
- Wat is discriminatie? | Mensenrechten voor jou | College voor de Rechten van de Mens. (n.d.). Retrieved August 18, 2024, from <https://www.mensenrechten.nl/mensenrechten-voor-jou/discriminatie-en-gelijke-behandeling/wat-is-discriminatie>
- Wiśniewski, J., & Biecek, P. (2022). Fairmodels: A flexible tool for bias detection, visualization, and mitigation in binary classification models. *The R Journal*, 14, 227–243. <https://doi.org/10.32614/RJ-2022-019>

Deel II

APPENDICES



APPENDIX A - FAIRNESS ANALYSESPROCES

A.1 ANALYSE VAN FAIRNESS

Het proces om tot een voorspelmodel te komen van retentie na 1 jaar en daarmee fairness te berekenen kent drie stappen: 1) het ontwikkelen van voorspelmodellen¹, 2) het bepalen van het belang van variabelen en 3) het berekenen van fairness².

A.1.1 *Stap 1 - Het ontwikkelen van voorspelmodellen*

De eerste stap is het prepareren van de data en het bouwen van een aantal voorspelmodellen.

- Nadat we de nodige procedures hebben doorlopen, verzamelen we eerst historische studiedata. Per eerstejaars student per studie en studiejaar – het cohort – is er een rij in de dataset met studiedata. We combineren data waar nodig en lossen eventuele fouten op.
- Rijen zonder uitkomstvariabele, retentie na 1 jaar, verwijderen we³.
- Dan lossen we rijen met ontbrekende gegevens op; voorspelmodellen kunnen daar namelijk niet mee omgaan. Het is van belang om missende waarden eerst goed te onderzoeken om te begrijpen waarom ze ontbreken. Zo hebben mbo studenten en studenten met een buitenlandse vooropleiding geen gemiddelde

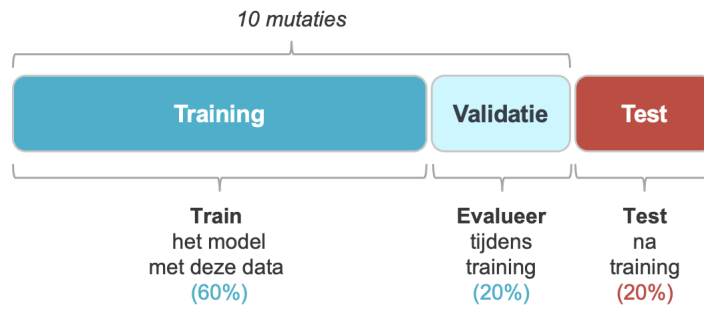
¹ Voor dit onderzoek hebben we gebruik gemaakt van het tidymodels package in R (Kuhn & Wickham, 2020).

² Voor het onderzoek naar variabelen en fairness hebben we gebruik gemaakt van de DALEX, DALEXtra en fairmodels packages in R (Biecek, 2018; Maksymiuk, Gosiewska, & Biecek, 2020; Wiśniewski & Biecek, 2022).

³ Dit kwam in ons onderzoek bij een opleiding voor die nog maar net gestart was.

cijfer van hun vooropleiding. We vullen rijen met ontbrekende gegevens op (imputatie). Per variabele die we opvullen, maken we een nieuwe variabele aan of er oorspronkelijke wel of geen waarde was (`VariabeleNaam_Missend`).

- Vervolgens onderzoeken we variabelen die sterk met elkaar correleren. Een voorbeeld is het gemiddelde cijfer voor het centraal eindexamen en het gemiddelde cijfer voor het schoolexamen. Van sterk gecorreleerde variabelen behouden we normaal gesproken de variabele die het meest voorspellend is, maar in dit onderzoek behouden we alle variabelen, zodat we naderhand de samenhang tussen deze variabelen nog kunnen begrijpen.
- Hierna verdelen we de data in drie delen: de trainingsdata, de testdata en de validatiedata. We splitsen de totale dataset bijvoorbeeld op in 60% om te trainen, 20% om te testen en 20% om het model dat we hebben ontwikkeld te valideren (Figuur A.1). In deze verdeling zorgen we ervoor dat de verhouding tussen studenten die wel of niet doorstuderen na 1 jaar overal gelijk is (stratificatie). Hoe goed het model presteert op de validatiedata zal de beoordeling van de kwaliteit van het model bepalen, zoals de accuraatheid.
- Nu maken we het voorspelmodel. Het is van belang om per opleiding een voorspelmodel te maken, omdat het voor elke opleiding uniek is welke studenten er instromen. De instroom hangt af van de vorm van de opleiding (associate degree of bachelor), de vorm van de opleiding (voltijd, deeltijd of duaal), de aard van de opleiding (het onderwerp) en voor geografische spreiding van de studenten de mate van concurrentie in de regio. Voor elke opleiding berekenen we twee soorten voorspelmodellen voor retentie na 1 jaar: een logistische regressie en een random forest model. We kiezen deze modellen omdat hiervan goed is uit te leggen welke variabelen er onderdeel van uit maken en wat de bijdrage per variabele is aan de voorspelling.
- Tot slot berekenen we per opleiding per model de kwaliteit met behulp van de Area Under the Curve (AUC/ROC). De AUC/ROC drukt uit in welke mate een voorspelmodel in staat is om positieve uitkomsten als positief te voorspellen en negatieve uitkomsten als negatief. In ons voorbeeld betekent dit dat het model goed voorspelt welke student er wel of niet doorstudeert na 1 jaar. Het model met de hoogste Area Under the Curve (AUC/ROC) vormt de basis voor het vervolg: de berekening van fairness.



© De Haagse Hogeschool, Lectoraat Learning Technology & Analytics, 2024

Figuur A.1: Verdeling van de dataset in drie delen: de trainingsdata, de testdata en de validatiedata

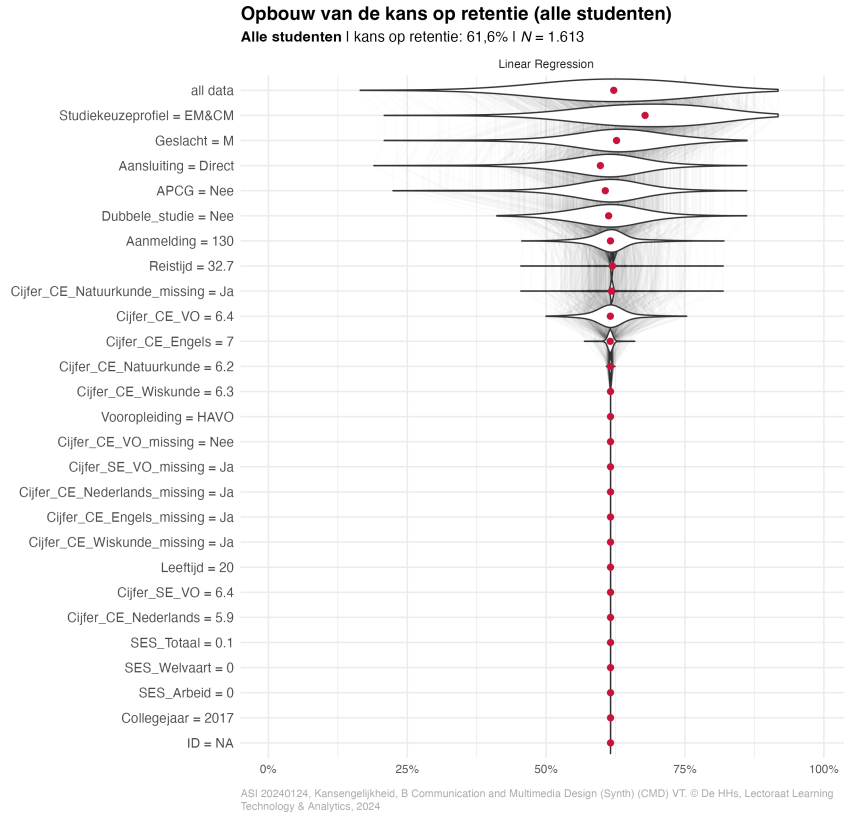
A.1.2 Stap 2 - Het bepalen van het belang van variabelen

Het tweede deel van de analyse bestaat uit een beter begrip van de variabelen in het voorspelmodel:

- Nadat we per opleiding twee modellen hebben ontwikkeld, gaan we verder met het model de hoogste mate van accuraatheid. Voor dat model berekenen we welke variabelen van belang zijn in de voorspelling en hoe ze bijdragen aan de kans op retentie na 1 jaar.
- Vervolgens passen we de modellen toe op de gemiddelde student van een opleiding. De gemiddelde student berekenen we door van de hele studentenpopulatie binnen de opleiding het gemiddelde te bereken bij continue variabelen en de meest voorkomende waarde bij categorische variabelen in de volgorde van belangrijkheid van de variabelen binnen het model. Zie Figuur A.2.
- Hierna herhalen we deze analyse, maar dan voor verschillende groepen. Hierbij onderzoeken we geslacht, hoogste vooropleiding en soort aansluiting⁴. Ook daar berekenen we per groep een gemiddelde student om er vervolgens het voorspelmodel op toe te passen. Zie Figuur A.3.
- Om de patronen die we daarin vinden beter te begrijpen passen we vervolganalyses toe om de rangorde van variabelen te beoordelen, de individuele effecten van variabelen beter te begrijpen en de mate van variantie binnen groepen te kunnen

⁴ Met soort aansluiting wordt bedoeld hoe de student na de hoogste vooropleiding is ingestroomd: is dit direct na de vooropleiding, na een tussenjaar, na een interne of externe switch, na een colloquium doctum c.q. 21+ toets, als 2e studie of is de aansluiting onbekend.

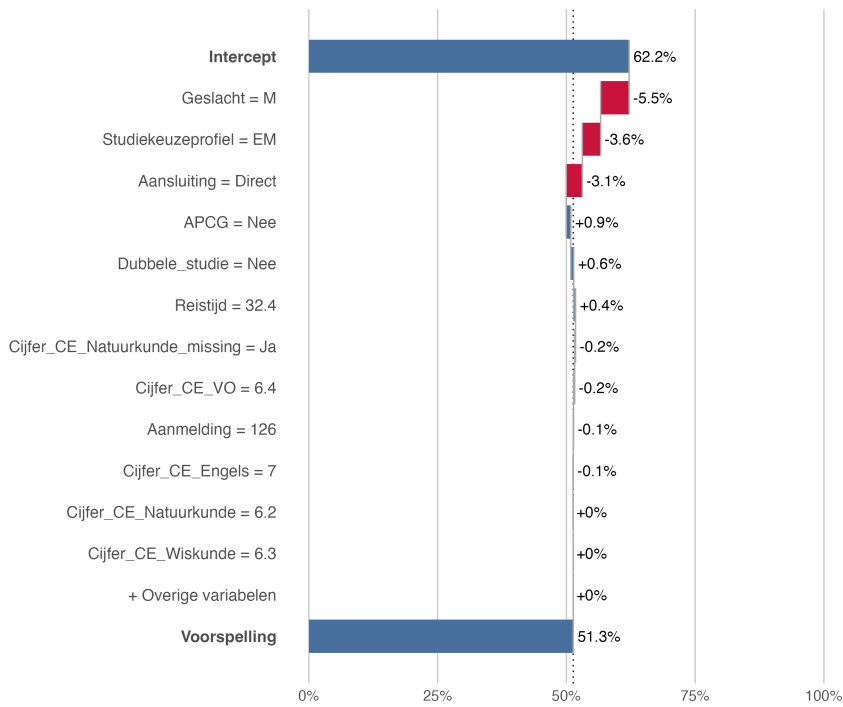
inschatten: een Shapley analyse, een Ceteris Paribus analyse (Figuur A.4) en een Partial Dependency analyse (Figuur A.5).



Figuur A.2: Toepassing van het model op een gemiddelde, fictieve student

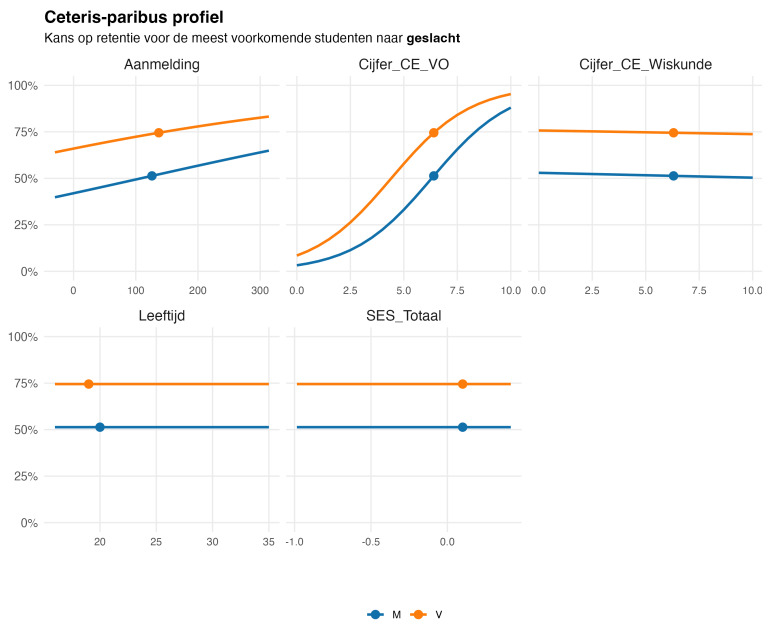
Opbouw van de kans op retentie naar geslacht

Geslacht: M | kans op retentie: 51,3% | N = 915 van 1.613 (56,7%)



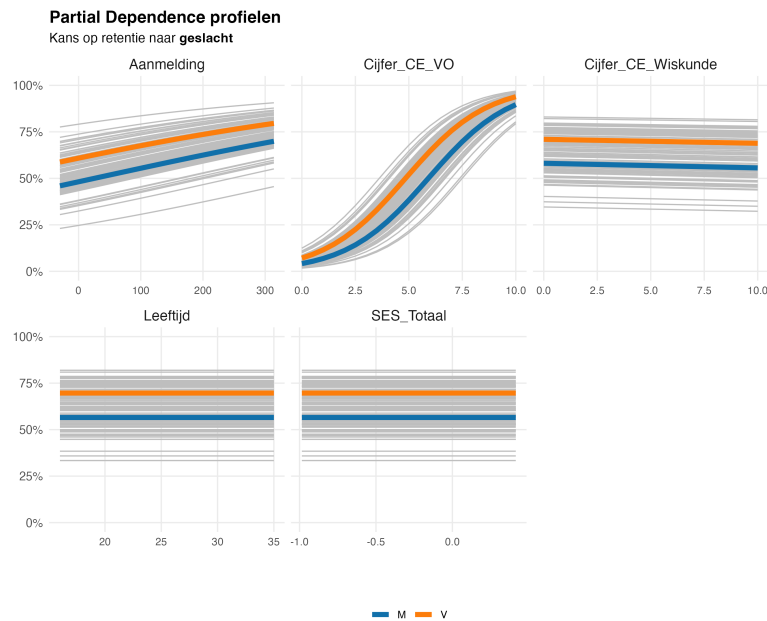
ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur A.3: Toepassing van het model op een gemiddelde, fictieve mannelijke student



ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur A.4: Ceteris Paribus Plot



ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur A.5: Partial Dependence Plot

A.1.3 Stap 3 - Het berekenen van fairness

Het derde en laatste deel van het proces is de berekening van fairness:

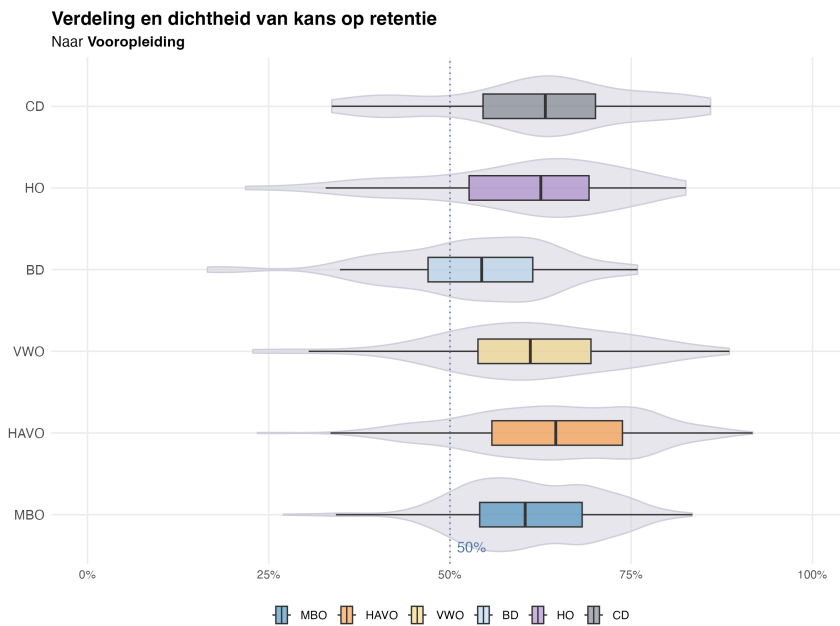
- Eerst analyseren we de kansverdeling binnen groepen; hiermee krijgen we een goed beeld per groep wat de verdeling is van de kansen in het voorspelmodel en kunnen die onderling ook goed vergelijken. Zie Figuur A.6.
- Vervolgens passen we het 4/5 criterium toe op een aantal prestatiewaarden, waaronder accuraatheid⁵. Als bevoorrechte groep hanteren we de meest voorkomende subgroep binnen een variabele⁶. Als er twee of meer waarden buiten de bandbreedte van het 4/5 criterium vallen is er sprake van een gebrek aan fairness voor de beschermde groep ten opzichte van de bevoorrechte groep. Dit kan negatief, maar ook juist positief uitvallen voor sommige groepen. Daarbij

⁵ In totaal zijn 5 ratio's beoordeeld: Accuracy Equality, Equal Opportunity, Predictive Equality, Predictive Parity en Statistical Parity. Zie Appendix A voor een toelichting op deze ratio's.

⁶ Stel dat er meer mannen dan vrouwen zijn in een opleiding, dan is Man de bevoorrechte groep; bij een gelijke stand wordt een van de twee gekozen.

berekenen we ook het 'totale verlies' per variabele. Hoe groter dit getal, des te groter de verschillen zijn binnen een opleiding. Zie Figuur A.7.

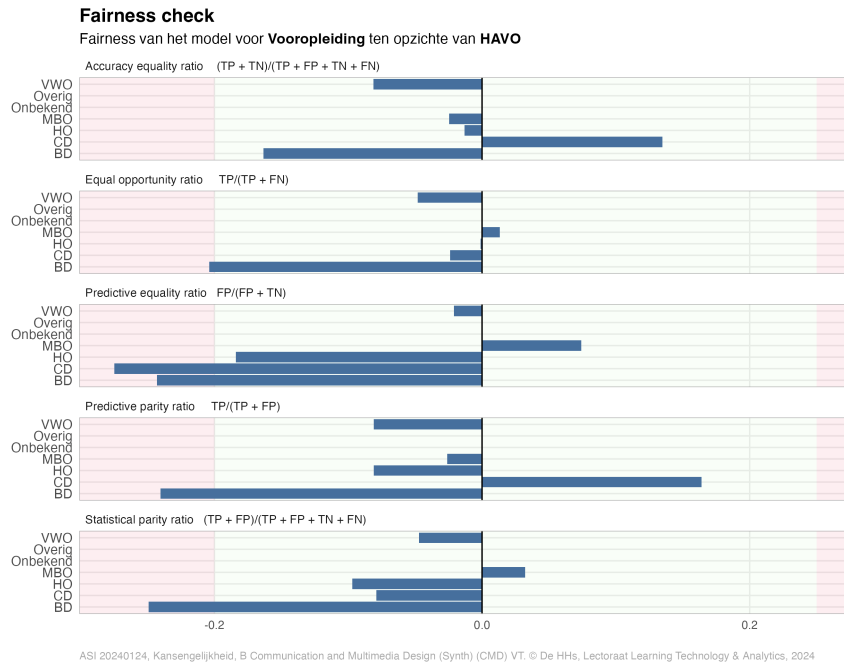
- Tot slot vatten we de bevindingen samen in een tabel om te zien waar er bias is en voor welke groepen. Daarbij maken we een onderscheid tussen groepen waarvan we voldoende data hebben om een uitspraak te doen of niet⁷. Zie Figuur A.8.



ASI 20240124, Kansengelijkheid, B Communication and Multimedia Design (Synth) (CMD) VT. © De HHs, Lectoraat Learning Technology & Analytics, 2024

Figuur A.6: Distributie van de kans op retentie na 1 jaar per hoogste vooropleiding

⁷ NB. Het kan zijn dat de bevoorrechte groep juist de groep is die last heeft van bias. Dit is vast te stellen als alle andere groepen een negatieve of positieve bias hebben



Figuur A.7: Fairness analyse per hoogste vooropleiding

Variabele	Groep	Bias	Geen Bias	Negatieve Bias	Positieve Bias
<i>Geslacht</i>	V	Nee	5	0	0
<i>Vooropleiding</i>	BD	Ja	1	3	1
	CD	Ja	0	0	5
	HO	Nee	5	0	0
	MBO	Ja	2	3	0
	VWO	Ja	3	0	2
<i>Aansluiting</i>	2e Studie	Ja	2	2	1
	Na CD	Ja	0	0	5
	Switch extern	Nee	4	0	1
	Switch intern	Nee	4	1	0
	Tussenjaar	Ja	2	0	3

Figuur A.8: Fairness conclusies per variabele

Deze procedure hebben we doorlopen voor alle associate degree en bachelor-opleidingen van De Haagse Hogeschool in elke variant. Per opleiding hebben we de procedure vervat in een zichzelf uitleggende website met per stap een eigen pagina inclusief

de onderliggende code. Vanaf vandaag wordt de broncode van deze methode door het lectoraat open source ter beschikking gesteld aan iedereen in het onderwijs die deze aanpak ook wil toepassen. De verdere verspreiding hiervan verloopt via het Npuls programma in samenwerking met de De Haagse Hogeschool.

B

APPENDIX B - FAIRNESS RATIO'S

Voor elke groep onderzoeken we 5 *ratio's*, ook wel maatstaven of metrieken genoemd. Deze *ratio's* zijn afgeleid van verhoudingen in de *confusion matrix*; ze geven inzicht in de mate van bias en kanselijkheid vanuit verschillende perspectieven van een prognosemodel.

		Voorspelde klasse		
		Positief	Negatief	
Werkelijke klasse	Positief	True Positive (TP)	False Negative (FN) Type II Error	② Sensitivity Equal opportunity ratio (EO) $\frac{TP}{(TP + FN)}$
	Negatief	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		④ Precision Predictive parity ratio (PPV) $\frac{TP}{(TP + FP)}$	Negative Predicted Value $\frac{TN}{(TN + FN)}$	① Accuracy Accuracy equality ratio (ACC) $\frac{TP + TN}{(TP + TN + FP + FN)}$
			③ Predictive opportunity ratio (PO) $\frac{FN}{(FN + TN)}$	⑤ Statistical parity ratio (STP) $\frac{TP + FN}{(TP + TN + FP + FN)}$

© De Haagse Hogeschool, Lectoraat Learning Technology & Analytics

Figuur B.1: Confusion Matrix met 5 *ratio's* voor fairness

B.1 ACCURACY EQUALITY

In welke mate voorspelt het model zowel de positieve als negatieve uitkomsten goed?

Deze maatstaf wordt gebruikt om te beoordelen of een model even accuraat is voor verschillende subgroepen binnen de dataset. Het vergelijkt de nauwkeurigheid van het model voor een beschermde groep (een minderheidsgroep) met de nauwkeurigheid voor een bevoorrechte groep (de meerderheidsgroep). Deze ratio wordt berekend als de verhouding tussen de nauwkeurigheid voor de beschermde groep en de nauwkeurigheid voor de bevoorrechte groep.

Formule

$$ACC = \frac{TP + TN}{TP + FP + TN + FN}$$

Definities

- *Nauwkeurigheid (Accuracy)*: Het percentage correcte voorspellingen van het model. Dit wordt berekend als het aantal juiste voorspellingen gedeeld door het totale aantal voorspellingen.

Interpretatie

- *Ratio = 1*: Het model is even accuraat voor beide groepen.
- *Ratio < 1*: Het model is minder accuraat voor de beschermde groep dan voor de bevoorrechte groep, wat wijst op mogelijke bias tegen de beschermde groep.
- *Ratio > 1*: Het model is meer accuraat voor de beschermde groep dan voor de bevoorrechte groep, wat ook op een vorm van bias kan wijzen, maar in het voordeel van de beschermde groep.

Voorbeeld

Stel dat we retentie voor havisten (bevoorrecht) en mbo-ers (beschermd) voorspellen en de nauwkeurigheid van het model voor havisten 80% is en voor mbo-ers 70%. De Accuracy Equality Ratio is dan 0,875 (70% / 80%). Dit betekent dat het model minder accuraat is voor de beschermde groep (mbo-ers) dan voor de bevoorrechte groep (havisten).

B.2 EQUAL OPPORTUNITY

In welke mate zijn de terecht positieve voorspellingen gelijk?

Deze maatstaf beoordeelt de gelijkheid van kansen die een model biedt aan verschillende subgroepen in termen van de 'True Positive Rate' (TPR). Het vergelijkt de kans dat een model correct een positieve uitkomst voorspelt voor een beschermde groep versus een bevoorrechte groep.

De True Positive Rate (TPR) wordt ook wel de *gevoeligheid* (*sensitivity* of *recall*) genoemd. De Equal Opportunity Ratio wordt berekend als het aantal true positives gedeeld door het totaal aantal werkelijke positives.

Formule

$$\text{TPR} = \frac{TP}{TP + FN}$$

Definities

- *TP*: True Positives (correcte voorspellingen van positieve uitkomsten)
- *FN*: False Negatives (werkelijke positieve uitkomsten die foutief als negatief zijn voorspeld)

Interpretatie

- *Ratio = 1*: Het model biedt gelijke kansen aan beide groepen in termen van het correct voorspellen van positieve uitkomsten.
- *Ratio < 1*: Het model biedt minder kansen aan de beschermde groep in vergelijking met de bevoorrechte groep, wat wijst op mogelijke bias tegen de beschermde groep.
- *Ratio > 1*: Het model biedt meer kansen aan de beschermde groep in vergelijking met de bevoorrechte groep, wat ook op een vorm van bias kan wijzen, maar in het voordeel van de beschermde groep.

Voorbeeld

Stel dat we retentie voor havisten (bevoorrecht) en mbo-ers (beschermd) voorspellen en de True Positive Rate voor havisten 70% is en voor mbo-ers 60%. De Equal Opportunity Ratio is dan 0,857 (60% / 70%). Dit betekent dat het model minder kans biedt aan de beschermde groep (mbo-ers) om correct positieve uitkomsten te voorspellen dan aan de bevoorrechte groep (havisten).

B.3 PREDICTIVE EQUALITY

In welke mate zijn de vals positieve voorspellingen gelijk?

Deze maatstaf beoordeelt de gelijkheid van het aantal False Positives (vals-positieven) tussen verschillende subgroepen. Het vergelijkt de False Positive Rate (FPR) voor een beschermde groep met de FPR voor een bevoorrechte groep.

De FPR is de verhouding van het aantal vals-positieve voorspellingen (False Positives, FP) ten opzichte van het totaal aantal werkelijke negatieve gevallen (True Negatives, TN en False Positives, FP). De Predictive Equality Ratio wordt berekend als de verhouding tussen de FPR voor de beschermde groep en de FPR voor de bevoorrechte groep.

Formule

$$FPR = \frac{FP}{FP + TN}$$

Definities

- *FP*: False Positives (foutieve voorspellingen van positieve uitkomsten)
- *TN*: True Negatives (correcte voorspellingen van negatieve uitkomsten)

Interpretatie

- *Ratio = 1*: Het model heeft een gelijke kans om False Positives te maken voor beide groepen.
- *Ratio < 1*: Het model heeft minder kans om False Positives te maken voor de beschermde groep dan voor de bevoorrechte groep, wat wijst op mogelijke bias ten nadele van de bevoorrechte groep.
- *Ratio > 1*: Het model heeft meer kans om False Positives te maken voor de beschermde groep dan voor de bevoorrechte groep, wat wijst op mogelijke bias ten nadele van de beschermde groep.

Voorbeeld

Stel dat we retentie voor havisten (bevoorrecht) en mbo-ers (beschermd) voorspellen en de False Positive Rate voor havisten 20% is en voor mbo-ers 30%. De Predictive Equality Ratio is dan 1,5 (30% / 20%). Dit betekent dat het model meer kans heeft om vals-positieve voorspellingen te maken voor de beschermde groep (mbo-ers) dan voor de bevoorrechte groep (havisten).

B.4 PREDICTIVE PARITY

In welke mate zijn de terecht positieve voorspellingen gelijk?

Deze maatstaf beoordeelt de gelijkheid van de nauwkeurigheid van de voorspelling tussen verschillende subgroepen door de positieve voorspellende waarde (Positive Predictive Value, PPV) van het model voor een beschermde groep te vergelijken met die voor een bevoorrechte groep. De PVV wordt ook wel de *precisie* genoemd en wordt berekend als het aantal true positives (TP) gedeeld door het totaal aantal voorspelde positives (TP en FP). Het is een maat voor de nauwkeurigheid van de positieve voorspellingen van het model.

Formule

$$PPV = \frac{TP}{TP + FP}$$

Definities

- *TP*: True Positives (correcte voorspellingen van positieve uitkomsten)
- *FP*: False Positives (foutieve voorspellingen van positieve uitkomsten)

Interpretatie

- *Ratio = 1*: Het model heeft een gelijke nauwkeurigheid in voorspellingen voor beide groepen.
- *Ratio < 1*: Het model is minder nauwkeurig in het voorspellen van positieve uitkomsten voor de beschermde groep dan voor de bevoorrechte groep, wat wijst op mogelijke bias tegen de beschermde groep.
- *Ratio > 1*: Het model is nauwkeuriger in het voorspellen van positieve uitkomsten voor de beschermde groep dan voor de bevoorrechte groep, wat ook op een vorm van bias kan wijzen, maar in het voordeel van de beschermde groep.

Voorbeeld

Stel dat we retentie voor havisten (bevoorrecht) en mbo-ers (beschermd) voorspellen en de Positive Predictive Value voor havisten 80% is en voor mbo-ers 70%. De Predictive Parity Ratio is dan 0,875 (70% / 80%). Dit betekent dat het model minder nauwkeurig is in het voorspellen van positieve uitkomsten voor de beschermde groep (mbo-ers) dan voor de bevoorrechte groep (havisten).

B.5 STATISTICAL PARITY

In welke mate zijn de positieve voorspellingen gelijk?

Deze maatstaf beoordeelt de gelijkheid van de positieve voorspellingen (ongeacht of ze correct zijn of niet) tussen verschillende subgroepen. Het vergelijkt de kans dat een model een positieve uitkomst voorspelt voor een beschermde groep met de kans dat het een positieve uitkomst voorspelt voor een bevoorrechte groep. De Statistical Parity Ratio wordt berekend als de verhouding tussen de kans op een positieve voorspelling voor de beschermde groep en de kans op een positieve voorspelling voor de bevoorrechte groep.

Formule

$$\text{SPR} = \frac{TP + FP}{TP + FP + TN + FN}$$

Definities

- *Positieve Voorspelling*: Een voorspelling waarin het model een positieve uitkomst voorspelt (bijv. aangenomen worden, krediet goedkeuring, etc.).

Interpretatie

- *Ratio = 1*: Het model voorspelt even vaak positieve uitkomsten voor beide groepen.
- *Ratio < 1*: Het model voorspelt minder vaak positieve uitkomsten voor de beschermde groep dan voor de bevoorrechte groep, wat wijst op mogelijke bias tegen de beschermde groep.
- *Ratio > 1*: Het model voorspelt vaker positieve uitkomsten voor de beschermde groep dan voor de bevoorrechte groep, wat kan wijzen op bias ten nadele van de bevoorrechte groep.

Voorbeeld

Stel dat we retentie voor havisten (bevoorrecht) en mbo-ers (beschermd) voorspellen en de kans op een positieve voorspelling voor havisten 80% is en voor mbo-ers 70%. De Statistical Parity Ratio is dan 0,875 (70% / 80%). Dit betekent dat het model minder vaak positieve uitkomsten voorspelt voor de beschermde groep (havisten) dan voor de bevoorrechte groep (mbo-ers).

TYPOGRAFIE

Dit document is opgemaakt met de typografische look-and-feel `classicthesis` ontwikkeld door André Miede en Ivo Pletikosić: <https://bitbucket.org/amiede/classicthesis/>.
Het omslag is ontwikkeld met behulp van een Latex-sjabloon van de TU Delft.

NO FAIRNESS WITHOUT AWARENESS

Toegepast onderzoek naar kansengelijkheid in het hoger onderwijs
Intreerede lectoraat Learning Technology & Analytics

In zijn intreerede bespreekt Dr. Theo C. Bakker kansengelijkheid in het hoger onderwijs en de rol van data-analyse in het bevorderen van sociale rechtvaardigheid. Vanuit zijn lectoraat Learning Technology & Analytics aan De Haagse Hogeschool voert Bakker onderzoek uit naar kansengelijkheid voor verschillende studentengroepen, waarbij hij voorspelmodellen en studiedata gebruikt om eventuele structurele ongelijkheden bloot te leggen. Hij benadrukt dat kansengelijkheid alleen kan worden opgelost door eerst het onderwijssysteem bewust te onderzoeken. *No Fairness without Awareness.*

Learning Analytics en bias – Learning analytics richt zich op het meten en analyseren van studentgegevens om onderwijs te verbeteren. Bakker onderscheidt hierin verschillende niveaus, zoals student analytics en institutional analytics, en focust op inclusion analytics, waarin gekeken wordt naar kansengelijkheid. Bias – systematische vooroordelen in data – kan vooroordelen in algoritmen versterken en zo kansengelijkheid veroorzaken. De onderzoeksmethode maakt gebruik van het 4/5-criterium, waarbij fairness in uitkomsten gemeten wordt door te kijken of de kansen voor de beschermde groep minstens 80% zijn van die van de bevoorrechte groep.

Onderzoeksaanpak – Bakker gebruikt machine learning om retentie na het eerste studiejaar te voorspellen en onderzoekt vervolgens verschillen tussen groepen studenten, zoals mbo- en vwo-studenten. Hij volgt drie stappen: (1) *Data voorbereiden en modellen bouwen*: Data worden opgesplitst en opgeschoond om accurate voorspelmodellen te maken. (2) *Variabelen analyseren*: Invloed van kenmerken op uitkomsten wordt beoordeeld voor verschillende groepen. (3) *Fairness berekenen*: Het 4/5-criterium wordt toegepast op *metrics* zoals accuraatheid en statistische gelijkheid om bias en ongelijkheden te identificeren.

Resultaten, aanbevelingen en vervolgonderzoek – Uit het onderzoek blijkt dat kansengelijkheid bij veel opleidingen ontbreekt, met name voor mannen en mbo-studenten, die een hogere kans op uitval hebben. Bakker adviseert sensitieve kenmerken zoals migratieachtergrond mee te nemen in analyses op basis van informed consent. Daarnaast pleit hij voor meer flexibiliteit in het beleid, geïnspireerd door maatregelen tijdens de coronacrisis, die een positief effect hadden op studiesucces.

Toekomstvisie – Bakker benadrukt dat niet elke ongelijkheid het gevolg is van discriminatie en roept op tot data-informed interventies om sociale rechtvaardigheid in het onderwijs te bevorderen. Zijn methode wordt open access beschikbaar gesteld, zodat ook andere instellingen deze kunnen toepassen en kansengelijkheid systematisch en bewust kunnen onderzoeken.