

Co-referaat Mark Kattenberg: het gebruik van *machine learning* bij de analyse van beleid

Welkom, fijn dat u tijd vrijgemaakt hebt voor deze bijzondere CPB lecture over het gebruik van *machine learning* bij het analyseren van beleid. We horen vaak dat *machine learning* uw wereld gaat veranderen, maar *machine learning* heeft uw wereld al verandert op manieren die we 15 jaar geleden voor onmogelijk hielden. Artsen gebruiken *machine learning* om op basis van foto's medische diagnoses te stellen. Mijn kinderen gebruiken *machine learning* om van foto's grappige filmpjes te maken, die ze vervolgens delen op social media. *Machine learning* verandert onze wereld, maar verandert het de analyse van de economie? Verandert *machine learning* het maken van beleid? En, zo ja, hoe gaan we dit proces dan begeleiden? Ik ga u hier de komende twintig minuten meer over vertellen.

Laat ik voordat ik begin mijzelf voorstellen. Ik ben Mark Kattenberg en heb mij in 2019 omgeschoold van 'typische CPB econoom' naar 'typische CPB data scientist'. Ik begon toen bij het CPB Data Science Team, een speciale onderzoeksgroep die is opgericht om uit te zoeken welke *machine learning* technieken het CPB kunnen helpen om betere economische analyses te maken. Wat mij trok was een mix van nieuwsgierigheid "Zijn die modellen dan echt zo goed?", maar ook van scepticisme "Het CPB is geen social media bedrijf. Wat kunnen wij er dan mee?". Ik zal dit aan u vertellen.

Hoe ga ik dat doen? Ik zal zo eerst uitleggen wat *machine learning* is, en hoe het verschilt van de analysetechnieken die we normaliter op het CPB gebruiken. Vervolgens zal ik uitleggen dat het gebruik van *machine learning* past in de CPB-traditie om de nieuwste onderzoekstechnieken uit de economische wetenschap toe te passen. Ik zal enkele voorbeelden geven hoe we *machine learning* hebben toegepast. En tot slot zal ik ingaan op het gebruik van *machine learning* bij het uitvoeren van beleid, en of dit wel kan als de data die *machine learning* gebruikt ongewenste patronen bevat, zoals discriminatie.

Wat is *machine learning*?

Maar laat ik allereerst uitleggen wat *machine learning* is, en hoe deze modellen verschillen van de technieken die ik als econoom voor mijn komst bij het Data Science Team gebruikte. Ik beperk me dan tot wat officieel 'supervised *machine learning*' heet, want anders sta ik hier morgen nog.

Normaal deed ik altijd onderzoek naar het effect van de ene variabele op de andere. Een klassieke vraag is dan bijvoorbeeld wat het effect is van onderwijs op het inkomen. Dus dan kijk je in de data en dan zie je dat hoogopgeleiden meer verdienen dan laagopgeleiden. En nu denken velen van jullie: "Ja, hoho, wacht even Mark, zo simpel is dat niet. Je moet voor heel veel dingen corrigeren. Voor iemands capaciteiten bijvoorbeeld. Als je dat doet dan vindt je een veel lager effect van scholing op het inkomen." En dat klopt natuurlijk. Maar dit kleine voorbeeld laat precies zien hoe we normaal onderzoek doen: we zijn geïnteresseerd in een effect, het effect dat we vinden hangt af van de variabelen die we in het model stoppen, en we besluiten op basis van theorie welke variabelen we toevoegen aan het model.

Machine learning werkt heel anders. *Machine learning* modellen zijn gericht op het voorspellen van een uitkomst, zoals het inkomen. En deze modellen zijn heel goed in het vinden van nuttige voorspellers, de spelden, in een hooiberg aan data. Dus je geeft ze eenvoudigweg een enorme bak

data om het inkomen te voorspellen, zoals het opleidingsniveau, het aantal gewerkte uren of de kleur van de ogen en het *machine learning* model selecteert zelf welke variabelen nuttig zijn en welke niet. Het aantal variabelen kan zelfs groter zijn dan het aantal observaties, dat zult u nog zien. Dus onthoud: de meeste *supervised machine learning* modellen bepalen *zelf en op basis van de data* welke informatie belangrijk is om te voorspellen. En dit is dus fundamenteel ander dan bij de reguliere technieken waar de onderzoeker dit doet.

Wat zijn dan de voor- en nadelen van *machine learning* ten opzichte van reguliere modellen? Het voordeel van deze technieken is dat ze een uitkomst goed voorspellen en dat ze variabelen combineren op basis van de data, niet op basis van theorie. In het beste geval kan dit ertoe leiden dat nieuwe verbanden in de data gevonden worden. Maar dit is wel lastig. Want een nadeel van *machine learning* is dat je in principe weinig leert over de wijze waarop de voorspelling tot stand komt, je leert dus weinig over het effect van variabelen op de uitkomst.

Wat ga ik verder doen? Ik zal kort betogen dat het gebruik van *machine learning* past in de CPB-traditie om *state-of-the-art* technieken te gebruiken. Vervolgens zal ik ingaan op de mogelijkheden die deze technieken bieden en ik eindig hoe eventuele risico's door 'verstandig gebruik' ondervangen kunnen worden.

Het gebruik van de nieuwste technieken past in de CPB traditie

Zoals gezegd past *machine learning* in de CPB-traditie om gebruik te maken van *state of the art* technieken. De tijd is beperkt, dus ik behandel de beginperiode van het CPB en de huidige tijd.

Ik begin natuurlijk met de eerste directeur van het CPB, Jan Tinbergen. Tinbergen was een van de pioniers van structurele macro-economische modellen en stond daarmee aan de wieg van het vakgebied econometrie. Hij ontving hiervoor in 1969 de nobelprijs voor de economie, samen met Ragnar Frisch. *Case made*, zou ik zo zeggen. Tinbergen schreef zijn modellen nog met potlood uit op ruitjespapier. Met de komst van computers kwam er niet alleen meer rekenkracht, maar werd er ook steeds meer data vastgelegd om te onderzoeken. In Nederland kunnen de gegevens uit bijvoorbeeld de belastingaangifte gecombineerd worden met opleiding of zorggebruik. Dit kan voor *alle* Nederlanders en uiteraard zijn deze gegevens geanonimiseerd en ze mogen ze alleen voor wetenschappelijk onderzoek gebruikt worden binnen een streng beveiligde digitale omgeving. Maar hierdoor hebben we bij het CPB dus al ruime ervaring met het werken met *big data*.

De laatste decennia is daar nog het gebruik van geavanceerde onderzoekstechnieken, *research designs*, bovenop gekomen om aannemelijk te maken dat een geschat effect causaal geïnterpreteerd kan worden. Guido Imbens heeft hier ook actief aan bijgedragen. Tijdens de CPB lecture van twee jaar geleden gaf de spreker van de hoofdlezing, Magne Mogstad, een mooi voorbeeld toen hij sprak over intergenerationele ongelijkheid. Hij vroeg zich toen af: hangt de kans dat iemand een uitkering heeft af van het gebruik van die uitkering door hun ouders. Maar uiteraard verschillen mensen die wel of geen uitkering hebben in heel veel opzichten van elkaar. En dat geldt ook als je ouders wel of geen uitkering hebben. Dit maakt het heel moeilijk om het effect van een uitkering van ouders goed te meten. De spreker keek daarom naar de uitkering voor arbeidsongeschiktheid in Noorwegen. Want in Noorwegen beslissen rechters of je deze uitkering krijgt, en de ene rechter is hierin coulanter dan de ander. En welke rechter jouw zaak behandelt is een kwestie van pech of geluk. De spreker gebruikte het feit of jij toevallig een coulante of strenge

rechter kreeg om het effect van de uitkering van ouders te isoleren. Een heel slim idee! En voor de mensen die er twee jaar geleden niet bij waren: Ja, als je ouders deze uitkering gebruiken dan heb je later een hogere kans om deze ook te gebruiken.¹ Veel CPB-publicaties passen dit soort *research designs* dan ook al toe op de *big data* uit de beveiligde microdata-omgeving van het CBS.

Het ligt dan ook voor de hand om ook die volgende stap te zetten en die andere grote ontwikkeling, *machine learning*, toe te passen. Ik zal dan ook nu gaan uitleggen hoe we deze technieken kunnen gebruiken om betere analyses van beleid te maken.

***Machine learning* maakt betere analyses van beleid mogelijk**

Ik denk dat *machine learning* op drie gebieden kansen biedt: *Machine learning* maakt het onder andere mogelijk om beter te *voorspellen*, om causale effecten nauwkeuriger te meten, en om nieuwe data te gebruiken. Ik zal dit toelichten.

Voorspellingen zijn nodig bij het maken van beleid. Het gezegde luidt niet voor niets: “regeren is vooruitzien”. Maar hé, dan wil je natuurlijk vooruitzien volgens de beste voorspelling! Maar hoe weten we nu hoe goed onze huidige voorspelling is en of het niet beter kan? Neem nu bijvoorbeeld de wereldhandel. Die voorspelt het CPB iedere maand met een BVAR op basis van 24 variabelen, die geselecteerd zijn op basis van economische theorie. Maar er zijn er natuurlijk duizenden variabelen die gebruikt kunnen worden! Zo hebben wij zelf eens elf duizend mogelijk relevante voorspellende variabelen geselecteerd. Handmatig gaan spitten in deze enorme berg data om variabelen te selecteren is ondoenlijk. Maar dat hoeft ook helemaal niet, want *machine learning* modellen bepalen zelf welke variabelen ze combineren om te voorspellen. En denk je in: we hebben maar zo’n *vierhonderd* maanden waarin we de wereldhandel observeren, terwijl we *elfduizend* verklarende variabelen hebben. Voor de empirisch ingestelden: we hebben bijna dertig keer zoveel verklarende variabelen als observaties. En toch kan zo’n model goed met deze data uit de voeten. Want wij vinden dat de *machine learning* modellen net zo goed voorspellen als ons model dat op basis van de theorie 24 variabelen gebruikt. Dat is best indrukwekkend, want de *machine learning* modellen gebruiken hiervoor geen economische theorie. Daarnaast blijkt hieruit dat het nu niet nodig is om een nieuw voorspelmodel te bouwen voor de wereldhandel. Maar dat hadden we niet geweten zonder deze test met *machine learning*.

Naast voorspellen kan *machine learning* de analyse van beleid beter maken door de effecten van beleid beter te schatten. Dat is belangrijk, want als we beleid maken willen we weten natuurlijk weten of deze het beoogde effect heeft gehad. Ik heb aan het begin van mijn verhaal verteld dat het effect dat we via onze standaardtechnieken vinden vaak afhangt van de controlevariabelen die we gebruiken. Dat is natuurlijk soms lastig: welke controlevariabelen moeten we opnemen en hoe dan? *Machine learning* maakt het mogelijk om dit te bepalen op basis van een brede set data en leidt daarbij tot een nuttige check of de gevonden effecten van beleid robuust zijn. Wel hangen ook de met *machine learning* gevonden effecten natuurlijk af van de gebruikte data.

¹ Zes tot twaalf procent punt afhankelijk van het moment waarop gemeten wordt. Zie Gordon B. Dahl, Andreas Ravndal Kostøl, Magne Mogstad, Family Welfare Cultures, *The Quarterly Journal of Economics*, Volume 129, Issue 4, November 2014, Pages 1711–1752, <https://doi.org/10.1093/qje/qju019>.

Dus *machine learning* kan gebruikt worden om te voorspellen, om effecten van beleid robuust te meten, maar ook om nieuwe databronnen te gebruiken. Data zoals tekst of foto's. Bij het data science team dachten we: dat gaan we ook eens proberen! Omdat het handig is om te beginnen met iets waar je heel vertrouwd mee bent stelden we ons de vraag: "Kunnen we op basis van *machine learning* onderzoeken waar economen over schrijven?" Dus op een blauwe maandag downloadden mijn collega's eens een half miljoen abstracts van RePEc beginnend in 1980 tot en met 2020. Dit is een enorm aantal, maar dit is duidelijk te veel om door een mens doorgelezen te worden. De data is publiek, dus mocht u het toch willen proberen, dan post Andrei zo wel even de links om de data te downloaden en analyseren.²

Deze data hebben ze schoongemaakt en vervolgens zoeken ze voor alle half miljoen papers welk ander paper er het meest op lijkt. En wat zie je dan? Papers die in een bepaald jaar gepubliceerd werden lijken het meest op andere papers die in hetzelfde jaar gepubliceerd werden. Daarmee lijkt het er dus op dat binnen de economische wetenschap onderwerpen maar kort in de mode zijn.

Vervolgens wilden we natuurlijk weten waar die papers over gaan. Andrei en Clemens hebben 'families' of clusters van papers gemaakt en binnen deze families hebben ze het paper geselecteerd dat het meest representatief is voor de familie. Waar gingen deze representatieve papers over? In 2009 gingen ze over de 'financiële crisis'. In 2013 gingen ze over de 'staatsschuldcrisis'. Ik laat het aan u om te gokken waar de meeste papers in 2020 over gingen.

Het analyseren van wetenschappelijke abstracts stelde ons in staat uit te zoeken wat trends zijn binnen economisch onderzoek. Inmiddels zijn we begonnen met het analyseren van vacatureteksten om inzicht te krijgen in gevraagde vaardigheden. Dit doen we overigens niet alleen, maar in een samenwerkingsverband met onder andere CBS, TNO en UWV.

Ik begon mijn verhaal door te laten zien dat het gebruik van *machine learning* past in de CPB traditie om de nieuwste onderzoekstechnieken te gebruiken. Vervolgens heb ik aangegeven dat deze technieken kunnen bijdragen aan betere voorspellingen, het robuuster meten van effecten van beleid dat ze het mogelijk maken om nieuwe data te gebruiken, zoals tekst. Ik eindig met enkele gedachten hoe we *machine learning* verstandig kunnen toepassen bij het uitvoeren van beleid.

² Code om de database te downloaden is te vinden op <https://github.com/andrei-dubovik/repec>. Trends in de data kunnen inzichtelijk gemaakt worden via <https://dubovik.eu/blog/repec>. De analyse door Andrei en Clemens die ik in deze speech bespreek gaat verder dan het tonen van de trends.

Kunnen we *machine learning* toepassen als de data ongewenste patronen bevat?

De laatste tijd is er terecht aandacht voor het risico dat het gebruik van *voorspellende algoritmes* door de overheid leidt tot oneerlijke uitkomsten. Voorspelmodellen gebruiken data om hun tot hun voorspellingen te komen en we weten dat er in die data ongewenste patronen kunnen zitten. Zo weten we dat er discriminatie voorkomt tegen minderheden op de arbeidsmarkt en de woningmarkt. We weten dat een hartaanval minder vaak herkend wordt bij vrouwen. Er is dan ook een vrees dat voorspelmodellen zoals *machine learning* deze ongewenste patronen versterken. In Nederland keert dit terug in het debat over Syri, dus of de overheid voorspelalgoritmen mag gebruik om fraude op te sporen.³ Momenteel wordt dit debat voornamelijk vanuit juridisch perspectief gevoerd: *mag* de overheid een algoritme gebruiken om bijstandsfraude op te sporen? *Mag* de overheid hiervoor verschillende databronnen combineren? *Mag* de overheid daarbij gebruik maken van variabelen als tweede nationaliteit? Dit juridisch perspectief is belangrijk, daar wil ik niets aan af doen. Wel wil ik hier een *economisch* perspectief tegenover plaatsen.

Als het gaat om de vraag ‘moeten we een voorspelmodel gebruiken’ stel ik als econoom twee vragen: “waarom gebruiken we een eigenlijk een voorspelmodel? En wat is het alternatief?”. Omdat we ergens onzeker over zijn, onzekerheid over wie we een medische behandeling moeten geven, onzekerheid wie we moeten ondersteunen of onderwerpen aan een fraudeonderzoek. Het wegnemen van deze onzekerheid heeft baten, omdat we middelen efficiënter kunnen inzetten. Dat is de reden dat we een voorspelmodel ontwikkeld hebben. Overigens hebben voorspelmodellen ook voordelen op het gebied van rechtvaardigheid. Ik heb u verteld over rechters in Noorwegen die besluiten of iemand in aanmerking komt voor een uitkering en dat het voor de toewijzing uitmaakte of je een strenge rechter kreeg. Als mensen beslissingen maken bestaat dus het risico dat gelijke gevallen ongelijk behandeld worden en een van mijn CPB-collega’s heeft aangetoond dat dit ook in Nederland gebeurt.⁴ Een voorspelmodel behandelt gelijke gevallen op gelijke wijze, het *kan* niet anders. Dus het verkleinen van onzekerheid en de gelijke behandeling van gelijke gevallen lijken me voordelen van het gebruik van *machine learning*.

Dan is de vraag: versterkt een voorspelmodel ongewenste patronen in de data? Wat in de discussie vooralsnog gemist wordt, is dat *machine learning* ongewenste patronen in de data absoluut kan versterken, maar deze óók kan tegengaan!⁵ Het hangt er maar juist van af hoe deze patronen

³ In Nederland komt dit tot uiting in het debat over het gebruik van voorspellende algoritmen door de overheid. Zie bijvoorbeeld deze berichten op de NOS. <https://nos.nl/artikel/2286848-overheid-gebruikt-op-grote-schaal-voorspellende-algoritmes-risico-op-discriminatie.html> en <https://nos.nl/artikel/2321704-anti-fraudesysteem-syri-moet-van-tafel-overheid-maakt-inbreuk-op-priveleven.html>.

⁴ In Nederland maken medewerkers verschillende keuzen bij het indelen van werklozen in arbeidsmarktintegratieprogramma’s, zie Jonneke Bolhaar, Nadine Ketel, Bas van der Klaauw, 2020, Caseworker's discretion and the effectiveness of welfare-to-work programs, *Journal of Public Economics*, Volume 183, <https://doi.org/10.1016/j.jpubeco.2019.104080>.

⁵ Rambachan en Roth onderzoeken de voorwaarden waaronder voorspellende algoritmen discriminatie versterken of verminderen. *Machine learning* gaat discriminatie tegen als er sprake is van selectieve labels bij de uitkomstvariabele op basis van taste-based discrimination, zie Rambachan, A., & Roth, J. (2019). Bias in, bias out? Evaluating the folk wisdom. *arXiv preprint arXiv:1909.08518*. <https://arxiv.org/abs/1909.08518>. Dat voorspellende algoritmen selectiviteit kunnen versterken of verminderen komt ook naar voren in de volgende papers:

ontstaan zijn. *Machine learning* modellen zorgen voor minder discriminatie als discriminatie in de data het gevolg is van voorkeuren van mensen (taste-based discrimination). We kunnen dus niet zeggen dat we nooit een voorspelmodel kunnen gebruiken als de data ongewenste patronen bevat. Dit hangt echt af van de vraag waarvoor we het voorspelmodel willen gebruiken en hoe het ongewenste patroon ontstaan is.

Maar laten we uitgaan van het negatieve scenario en aannemen dat het gebruik van een voorspelmodel leidt tot meer ongewenste patronen in de samenleving, zoals meer discriminatie tegen minderheden. Dat is onrechtvaardig en brengt in die zin hoge kosten met zich mee. Maar, onzekerheid en daarmee brengt het niet gebruiken van zo'n model óók kosten met zich mee. Dat is óók ongewenst. Hier zien we dus een klassieke uitruil tussen efficiëntie en rechtvaardigheid.⁶

En deze uitruil kunnen we expliciet maken: Gemeenten zijn verplicht om onderzoek te doen naar fraude en stel dat zij in dertig procent van de gevallen huishoudens met een tweede nationaliteit onderzoeken. Ik verzin zomaar een percentage, overigens. Waarom dan geen algoritme gebruiken om te voorspellen wie fraude zou kunnen plegen, terwijl we daarnaast óók nog opleggen dat maximaal dertig procent van de onderzochte huishoudens een tweede nationaliteit mag hebben? Op deze wijze benutten we dat we het risico op fraude waarschijnlijk beter inschatten, een winst in efficiëntie, terwijl we niet onrechtvaardiger zijn dan wanneer we geen voorspelmodel gebruiken.

Natuurlijk zitten er haken en ogen aan deze oplossing. Discriminatie kan bijvoorbeeld op meerdere gronden plaatsvinden dan alleen het hebben van een tweede nationaliteit. En de discussie hoe hoog het percentage moet zijn is lastige, ethische discussie. Maar de uitruil tussen efficiëntie en rechtvaardigheid werpt een ander licht op de vraag of we voorspelmodellen moeten gebruiken en hoe we dit rechtvaardig kunnen doen.⁷

Dan tot slot: betekent dit dat we in iedere situatie een voorspelmodel kunnen gebruiken? Nee, ik denk dat het heel belangrijk is om goed na te denken over de uitkomst die je wilt voorspellen. *Machine learning* modellen voorspellen precies datgene wat je ze vraagt te voorspellen. Dus een

Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S., 2020, An Economic Perspective on Algorithmic Fairness. *AEA Papers and Proceedings*, Vol. 110, pp. 91-95, <https://doi.org/10.1257/pandp.20201036>;
Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S., 2021, An Economic Perspective on Algorithmic Fairness. *NBER working paper 27111*, <https://doi.org/10.3386/w27111>;
Mullainathan, S., & Obermeyer, Z. (2017). Does *machine learning* automate moral hazard and error?. *American Economic Review*, 107(5), 476-80, <https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.p20171084> en
Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The quarterly journal of economics*, 133(1), 237-293. <https://doi.org/10.1093/qje/qjx032>.

⁶ Deze uitruil komt expliciet aan de orde in een serie artikelen. Zie Kleinberg, J., Ludwig, J., Mullainathan, S., Rambachan, A., 2018, Algorithmic fairness, *Aea papers and proceedings*, 108, pp. 22-27, <https://doi.org/10.1257/pandp.20181018>;

Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S., 2020, An Economic Perspective on Algorithmic Fairness. *AEA Papers and Proceedings*, Vol. 110, pp. 91-95, <https://doi.org/10.1257/pandp.20201036> en
Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S., 2021, An Economic Perspective on Algorithmic Fairness. *NBER working paper 27111*, <https://doi.org/10.3386/w27111>.

⁷ De uitruil tussen efficiëntie en rechtvaardigheid werpt ook een ander licht op de vraag of een voorspelmodel toegang moet hebben tot variabelen zoals 'tweede nationaliteit'. Ik weet dat het juridisch niet mag, maar deze uitruil impliceert ook dat het vanuit welvaartseconomisch oogpunt optimaal is om vooraf zoveel mogelijk variabelen mee te nemen en achteraf te corrigeren voor rechtvaardigheid. Zie (wederom) deze artikelen: Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S., 2020, An Economic Perspective on Algorithmic Fairness. *AEA Papers and Proceedings*, Vol. 110, pp. 91-95, <https://doi.org/10.1257/pandp.20201036>.

voorwaarde voor verstandig gebruik, is dat we datgene wat we willen verklaren goed kunnen meten.⁸ Dit is nog lastiger dan het lijkt, want de ene variabele kan leiden tot ongewenste uitkomsten en de andere niet. Dit inzicht is nog onderbelicht in de discussie, die vooral draait om het gebruik van de verklarende variabelen. Waarom is nadenken over wat je wilt voorspellen zo belangrijk? Ik geef een voorbeeld: In de Verenigde Staten hebben onderzoekers eens een bestaand voorspelmodel gebruikt dat tot doel had om mensen te selecteren met een slechte gezondheid, want die kwamen dan in aanmerking voor een extra behandeling.⁹ Nu kan je gezondheid op meerdere manieren meten, dit kan via zorguitgaven of via het aantal chronische aandoeningen, beide manieren gebruiken we vaak als economen. De onderzoekers vonden dat het model minderheden benadeelde als het model gevraagd werd om zorgkosten te voorspellen. Maar dit gebeurde niet als het model gevraagd werd om het aantal chronische aandoeningen te voorspellen. Maar dit is hetzelfde model, hetzelfde optimalisatie-algoritme, getraind op dezelfde data. Oh, over de data, dit model wist niet wie tot een minderheid behoorde. Hier zien we een voorbeeld waarin een algoritme een ongewenst patroon in de data kan versterken, afhankelijk van *de uitkomst* die voorspeld wordt. Het is dus belangrijk om na te denken of je datgene wat je wilt voorspellen goed meet. Bij twijfel, wil je misschien geen voorspelmodel gebruiken.

Dus in dit laatste deel ben ik kort ingegaan op de vraag hoe we als econoom naar het gebruik van voorspelmodellen kunnen kijken. Ik heb uitgelegd dat voorspelmodellen gebruikt worden om onzekerheid weg te nemen. Ik heb uitgelegd dat het niet waar is dat *machine learning* altijd ongewenste patronen in de data versterkt. Ik heb uitgelegd dat het al dan niet gebruiken van een voorspellend algoritme zorgt voor een uitruil tussen efficiëntie en rechtvaardigheid en dat we kunnen voorkomen dat *machine learning* discriminatie in de data versterkt door achteraf restricties op te leggen aan het model. Tot slot heb ik uitgelegd dat het gebruik van *machine learning* wel vereist dat we hetgeen we willen verklaren goed kunnen meten.

Ik rond af. Ik heb laten zien dat het gebruik van *machine learning* past in de CPB-traditie om de nieuwste technieken te gebruiken en ik heb aangegeven hoe ze onze analyse van beleid kunnen verbeteren. Tot slot heb ik de olifant in de kamer bij zijn slurf gegrepen en ben ik ingegaan op het verstandig gebruik van *machine learning* bij het uitvoeren van beleid. En dat is belangrijk: Of we het nu willen of niet: *machine learning* en *big data* hebben onze wereld al veranderd en ze hebben de manier waarop we beleid maken en analyseren veranderd. En dit proces zal blijven doorgaan, deze ontwikkeling is eenvoudigweg niet te stoppen. En dat hoeft ook niet, mits we het proces goed begeleiden. Ik hoop met deze lecture nieuwe inzichten te hebben gegeven hoe we dat kunnen doen.

⁸ Zie bijvoorbeeld Mullainathan, S., & Obermeyer, Z. (2017). Does *machine learning* automate moral hazard and error?. *American Economic Review*, 107(5), 476-80, <https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.p20171084>.

Ik beperk mij hier tot een technische voorwaarde. Er zijn meer voorwaarden die ervoor kunnen zorgen dat voorspellende algoritmen verstandig gebruikt kunnen worden. De Algemene Rekenkamer heeft hier een rapport over uitgebracht. Zie Algemene Rekenkamer, *Aandacht voor algoritmes*, rapport, 26-01-2021, zie <https://www.rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmes>.

⁹ Het onderzoek wordt beschreven in Ziad Obermeyer, Brian Powers, Christine Vogeli, Sendhil Mullainathan, 2019, Dissecting racial bias in an algorithm used to manage the health of populations, *Science*, vol 366(6464), pp 447 – 453, DOI: <https://doi.org/10.1126/science.aax2342>.