



Universiteit
Leiden
The Netherlands

Bayesian learning: Challenges, limitations and pragmatics

Heide, R. de

Citation

Heide, R. de. (2021, January 26). *Bayesian learning: Challenges, limitations and pragmatics*. Retrieved from <https://hdl.handle.net/1887/3134738>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3134738>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3134738> holds various files of this Leiden University dissertation.

Author: Heide, R. de

Title: Bayesian learning: Challenges, limitations and pragmatics

Issue Date: 2021-01-26

Samenvatting

Dit proefschrift gaat over het leren van data op een Bayesiaanse wijze. *Statistiek en machine learning* gaan over de vraag hoe mensen en computers kunnen leren van data. Bayesiaanse methodes worden in deze vakgebieden veel gebruikt, echter, ze hebben bepaalde beperkingen en interpretatieproblemen die niet altijd worden onderkend. In twee hoofdstukken van dit proefschrift onderzoeken we een dergelijke beperking en omzeilen we deze door een verruiming van het standaardkader van de Bayesiaanse methode. In twee andere hoofdstukken nemen we door hoe verschillende filosofische interpretaties van het Bayesianisme wiskundige definities en stellingen beïnvloeden, en hoe dat zijn uitwerking heeft op de praktische toepassing van Bayesiaanse methodes. In de overige twee hoofdstukken passen we zelf Bayesiaanse methodes toe op een *pragmatische* wijze: enkel als *werktuig* voor een interessant statistisch probleem, een probleem dat ook op een niet-Bayesiaanse manier had kunnen worden aangepakt.

Leren Als een onderzoeker iets wil leren over een onbekend proces, vindt er een interactie plaats tussen haar en de *data* die door het proces zijn voortgebracht. De taak van de onderzoeker is *inductie*: een manier van redeneren waarbij er op grond van waarnemingen tot een algemene regel — een generalisatie — wordt gekomen. De onderzoeker begint met enkele veronderstellingen over het onbekende proces, omdat zonder deze *voorkennis* de datapunten op iedere mogelijke manier zouden kunnen samenhangen en het onmogelijk is tot een generalisatie te komen. Daarnaast bestaat er een verzameling van *hypotheses* die de onderzoeker kan opstellen of onderzoeken: algemene beschrijvingen van het onbekende proces. In de context van dit proefschrift, statistiek en machine learning, beschouwen we hypotheses die kunnen worden uitgedrukt als een waarschijnlijkheidsverdeling over een uitkomstenruimte, en deze noemen we *statistische hypotheses*. Een verzameling statistische hypotheses vormt een (statistisch) *model*. Een model is een wiskundige weergave van de voorkennis.

Bayesianisme Naast een model en de data hebben we een laatste ingrediënt nodig voor inductie: een methode. Het hoofdthema van dit proefschrift is de *Bayesiaanse* methode. In essentie is dit een methode die niet alleen waarschijnlijkheidsverdelingen over de data hanteert, maar ook over de statistische hypotheses. De onderzoeker begint met het specificeren van een *prior*, een waarschijnlijkheidsverdeling die haar onzekerheid over de statistische hypotheses uitdrukt, voordat ze heeft kennis genomen van de data. Na waarneming van de data, wordt met *de stelling van Bayes* een *posterior* berekend: een conditionele waarschijnlijkheidsverdeling

over de statistische hypotheses gegeven de data.

Bayesianisme is een term die verwijst naar een verzameling aan deze methode verwante ideeën in verschillende takken van de wetenschap. Echter, *het* Bayesianisme bestaat niet: er zijn verschillende stromingen, die er bijvoorbeeld verschillende theoriën op na houden over hoe de priors tot stand komen. Twee noemenswaardige, invloedrijke stromingen zijn het *subjectivisme* en het *objectivisme*. In de tweede helft van dit proefschrift staat een derde stroming centraal: het *pragmatisme*: onderzoekers die de Bayesiaanse methode niet uit filosofische overtuigingen bezigen, maar enkel vanwege haar nuttige eigenschappen of andere praktische beweegredenen. Discussies over de grondslagen van het Bayesianisme worden vaak in de filosofie gevoerd; welke stroming men aanhangt heeft nochtans consequenties voor de (statistische) praktijk: welke priors men kiest, welke wiskundige definities men formuleert en welke stellingen men poneert, hangt hier vanaf.

Misspecificatie van het model Zoals hierboven beschreven, begint de onderzoeker met het specificeren van een model en het toekennen van priorwaarschijnlijkheden aan zijn elementen. Als het ware datagenererende proces onderdeel is van het model, en niet uitgesloten wordt door de prior, is *consistentie* gegarandeerd: naar mate we meer en meer data verkrijgen, valt de onderzoekers posterior meer en meer samen met de ware verdeling. Niettemin kan het voorkomen dat het model *gemisspecificeerd* is: het ware datagenererende proces is geen onderdeel van het model (of heeft prior nul toegekend gekregen). Dit kan op verschillende manieren problematisch zijn en in dit proefschrift wordt de Bayesiaanse methode op twee verschillende manieren uitgebreid om twee van deze problemen te boven te komen.

Ten eerste kan het gebeuren dat de onderzoeker tijdens het leerproces een nieuwe hypothese bedenkt en deze wil toevoegen aan het model. In het standaardkader van de Bayesiaanse methode is dit in principe niet mogelijk: de onderzoeker moet de reeds verkregen data weggooien en opnieuw beginnen met het toekennen van priorwaarschijnlijkheden aan de elementen van het nieuwe, grotere model. In hoofdstuk 2 bestuderen we een *ruimdenkende Bayesiaanse logica*, die het dynamisch bijvoegen van nieuwe hypotheses tijdens het leerproces mogelijk maakt.

Ten tweede kan het gebeuren dat we willen dat de Bayesiaanse posterior samenvalt met het *beste* element in het model, in plaats van met de ware verdeling die buiten het model ligt. In hoofdstuk 6 laten we zien hoe dit kan mislukken met de Bayesiaanse standaardmethode. Vervolgens verrichten we een aanpassing aan de stelling van Bayes: de aannemelijkheidsverdeling wordt tot een macht verheven, en we noemen dit de *gegeneraliseerde Bayesiaanse methode*. Indien deze macht gevoelig wordt gekozen, lost dit het probleem op, en valt de gegeneraliseerde Bayesiaanse posterior na vergaring van data samen met het beste element in het model, ondanks de modelmisspecificatie.

Optioneel stoppen met de Bayes-factor-hypothesetoets De *Bayes factor* is een Bayesiaanse methodiek voor hypothesetoetsen. In hoofdstuk 3 en 4 bestuderen we *optioneel stoppen*. In formeel betekent dit ‘tijdens het leerproces naar de tussenresultaten kijken om te beslissen of er meer datapunten vergaard moeten worden’. Verschillende auteurs beweren dat Bayesiaanse methodes *bestand zijn tegen optioneel stoppen*, maar het blijkt onduidelijk te zijn wat dat precies betekent. In hoofdstuk 4 geven we drie verschillende wiskundige definities van deze uitspraak.

In hoofdstuk 3 en 4 zetten we uiteen hoe het aanhangen van een van de stromingen van het Bayesianisme invloed heeft op welke beweringen men kan doen in de praktijk. In hoofdstuk 3 laten wij bijvoorbeeld zien dat sommige beweringen over optioneel stoppen met Bayes factors alleen betekenis hebben als ze vanuit een puur subjectieve invalshoek worden gedaan, desalniettemin worden deze beweringen vaak gedaan als zouden ze ook gelden voor een pragmatisch Bayesiaanse benadering.

Een nieuwe theorie voor hypothesetoetsen In hoofdstuk 5 presenteren we een nieuwe theorie voor hypothesetoetsen. Deze theorie draait om het concept genaamd ‘ ϵ -variabele’ of ‘ ϵ -waarde’, een stochast die de mate van bewijs tegen de nulhypothese aanduidt en die in de toekomst hopelijk de p -waarde zal vervangen in de toegepaste statistiek. Tevens introduceren we een optimaliteitscriterium voor de constructie van ϵ -variabelen, genaamd *GROW*, wat een acroniem is voor het Engelse *Growth-Optimal in Worst Case*. Het blijkt dat de *GROW* ϵ -variabele een Bayesiaanse interpretatie kent, zij het met een geheel ander soort priors dan priors die in de huidige Bayesiaanse praktijk worden gebruikt.

Identificatie van de beste waarschijnlijkheidsverdeling In hoofdstuk 7 bestuderen we een Bayesiaanse manier om uit een verzameling waarschijnlijkheidsverdelingen degene met de hoogste verwachtingswaarde te onderscheiden. We kunnen aan de verschillende verdelingen, die ook wel *armen* worden genoemd, een prior toekennen die de waarschijnlijkheid uitdrukt dat deze verdeling de hoogste verwachtingswaarde heeft. Vervolgens stellen we een regel op om op ieder tijdstip een arm te kiezen waarvan we een observatie willen ontvangen. Nadien berekenen we de posteriorwaarschijnlijkheid dat deze arm de hoogste verwachtingswaarde heeft. In hoofdstuk 7 bewijzen we asymptotische frequentistische garanties voor deze Bayesiaanse strategie.