



Universiteit
Leiden
The Netherlands

The many faces of online learning

Hoeven, D. van der

Citation

Hoeven, D. van der. (2021, March 4). *The many faces of online learning*. Retrieved from <https://hdl.handle.net/1887/3147345>

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/1887/3147345>

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

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3147345> holds various files of this Leiden University dissertation.

Author: Hoeven, D. van der

Title: The many faces of online learning

Issue Date: 2021-03-4

Samenvatting

Online Learning is een fundamentele machine learning setting waarin een leerder sequentieel voorspellingen moet doen gegeven (partiële) informatie over voorgaande correcte voorspellingen en mogelijk extra informatie. Over de omgeving van de leerder wordt vaak aangenomen dat het een vijandige omgeving is die de taak van de leerder, zo min mogelijk verlies lijden, zo moeilijk mogelijk maakt. Desalniettemin zijn er in de afgelopen drie decennia veel Online Learning algoritmes ontwikkeld die bevredigende garanties bieden in verschillende settings. De garanties van Online Learning algoritmes gaan over spijt: het verschil tussen het cumulatieve verlies van de leerder en het cumulatieve verlies van een offline optimizer van het cumulatieve verlies, waar de offline optimizer ook wel bekend staat als de vergelijker. In dit proefschrift presenteren wij verschillende nieuwe inzichten in verschillende settings van Online Learning. Vandaar ook de titel van het proefschrift: de vele gezichten van Online Learning.

In Hoofdstuk 2 bestuderen we een van de fundamenteelste algoritmes in Online Learning: Exponential Weights. We laten zien hoe Exponential Weights moet worden afgestemd zodat het in verschillende settings kan worden toegepast en we laten tevens zien hoe met specifieke keuzes voor de parameters verschillende belangrijke algoritmes worden teruggevonden als een speciaal geval van Exponential Weights. Dit inzicht leidt tot een gecentraliseerd begrip van verscheidene algoritmes in Online Learning en verenigt de analyse van deze algoritmes.

Een belangrijk onderscheid in Online Learning is het verschil tussen de volledige-informatie en bandit settings. In de volledige-informatie setting onthult de omgeving de gehele verliesfunctie, maar in de meer uitdagende bandit setting onthult de omgeving enkel partiële informatie. Een belangrijke eigenschap van veel algoritmes in zowel de volledige-informatie setting als de bandit setting is dat ze gepaste grenzen op de spijt garanderen, zelfs in een vijandige omgeving. Echter worden deze algoritmes dusdanig afgesteld dat ze enkel met een vijandige omgeving om kunnen gaan en niet meer goedaardige omgevingen kunnen uitbuiten. Een van de hoofdonderwerpen in dit proefschrift is hoe men algoritmes kan ontwerpen die zowel in een vijandige omgeving als in een goedaardige omgeving bevredigende garanties bieden, zonder dat van tevoren bekend is in wat voor omgeving de leerder zich bevindt. Omdat deze algoritmes zich aanpassen aan de omgeving staan dit

soort algoritmes bekend als adaptieve algoritmes. In Hoofdstuk 2 laten we zien hoe we verschillende adaptieve algoritmes terugvinden als een speciaal geval van Exponential Weights. In Hoofdstukken 3 en 4 bestuderen we een speciaal soort adaptief algoritme, namelijk vergelijker-adaptieve algoritmes. De spijt grens van een vergelijker-adaptief algoritme hangt af van de norm van de vergelijker, wat in sommige gevallen tot minder spijt kan leiden vergeleken met standaard algoritmes. In Hoofdstuk 3 laten we zien hoe we vergelijker-adaptieve algoritmes kunnen aanpassen zodat ze adaptief zijn aan onbekende ruis. Dit is nuttig wanneer mensen willen kiezen hoeveel privacy zij hebben maar dit niet willen laten weten. In Hoofdstuk 4 ontwikkelen wij de eerste vergelijker-adaptieve algoritmes in de Bandit Convex Optimisation setting. In Hoofdstuk 5 presenteren wij MetaGrad, dat zich kan aanpassen aan een grote klasse van functies.

Omdat de leerder in elke ronde zijn voorspellingen update is een belangrijke eigenschap van een Online Learning algoritme de looptijd. De per ronde looptijd wordt vaak als te hoog gezien wanneer de updates meer tijd in beslag nemen dan kwadratisch in de dimensie van het probleem. Hierdoor is er een aanzienlijke hoeveelheid werk verzet door verschillende auteurs om de looptijd van Online Learning algoritmes te verminderen. Ook in dit proefschrift is er aandacht besteed aan het versnellen van Online Learning algoritmes. In Hoofdstuk 6 presenteren we een nieuw algoritme voor de Online Multiclass Classification setting die vaak vergelijkbare of betere garanties biedt op de spijt dan tragere algoritmes. In deze setting moet de leerder in elke ronde een label voorspellen gegeven een d -dimensionale feature vector die mogelijk extra informatie bevat. In de Online Multiclass Classification setting lijdt de leerder de één-nul verliesfunctie. De één-nul verliesfunctie is één wanneer de voorspelling van de leerder incorrect is en nul wanneer de voorspelling van de leerder correct is. De benchmark in de Online Multiclass Classification setting is een convexe surrogaat verliesfunctie. Deze surrogaat verliesfunctie is een bovengrens op de één-nul verliesfunctie. Het doel van de leerder in de Online Multiclass Classification setting is de surrogaat spijt te minimaliseren: het verschil tussen de som van de één-nul verliesfuncties en het offline minimum van de som van de surrogaat verliesfuncties. Voorgaande algoritmes in de Online Multiclass Classification setting bouwden vaak voort op tweede-orde algoritmes om lage spijt te garanderen. Tweede-orde algoritmes houden een d bij d matrix bij die elke ronde wordt geupdate, wat de per ronde looptijd minstens d^2 maakt. Wij introduceren een nieuw algoritme genaamd GAPTRON dat een per ronde looptijd heeft van $O(d)$. GAPTRON heeft vaak een vergelijkbare of zelfs betere garantie op de spijt dan tragere algoritmes. Bijvoorbeeld, in de Bandit Online Multiclass Classification setting is de surrogaat spijt bovengrens van GAPTRON een factor $\sqrt{d}$ kleiner dan de surrogaat spijt bovengrens van langzamere algoritmes. We behalen deze resultaten

door het gat tussen de één-nul verliesfunctie en de surrogaat verliesfunctie uit te buiten. Door deze nieuwe aanpak kan de leerder gebruik maken van eerste-orde algoritmes om zijn voorspellingen te updaten en tegelijkertijd kleine surrogaat spijt te garanderen.

Verdere verbeteringen van looptijd worden gemaakt in Hoofdstuk 5. In Hoofdstuk 5 laten we zien hoe de looptijd van MetaGrad kan worden verlaagd met behulp van sketching methoden. In Hoofdstuk 7 bestuderen we de online portfolio selectie setting. Het optimale algoritme voor de online portfolio selectie setting is een versie van Exponential Weights. Helaas is de looptijd van deze specifieke versie van Exponential Weights te groot om als praktisch te worden beschouwd. Voor de online portfolio selectie setting zijn veel andere algoritmes overwogen, maar allemaal hebben ze tekortkomingen. We stellen een open probleem waarin we vragen om een snel en optimaal algoritme. Wij geven een gedeeltelijke analyse van wat wij denken dat een snel en optimaal algoritme is. Hierbij laten we zien dat het algoritme wat wij voorstellen inderdaad de optimale bovengrens op de spijt behaalt in specifieke gevallen.