Warszawa, 21.11.2023

UNIWERSYTET WARSZAWSKI
Wydział Matematyki, Informatyki i Mechaniki
Instytut Informatyki

Recenzja rozprawy doktorskiej magistra Macieja Wołczyka pod tytułem *Adaptiveness in Deep Learning Models*

Artykuły na których bazuje rozprawa

[1] Wołczyk, Maciej, Michał Zając, Razvan Pascanu, Łukasz Kuciński, Piotr Miłoś "Continual World: A robotic benchmark for continual reinforcement learning." *Advances in Neural Information Processing Systems* (NeurIPS), pp. 28496-28510, 2021.

[2] Wołczyk, Maciej, Michał Zając, Razvan Pascanu, Łukasz Kuciński, Piotr Miłoś, "Disentangling Transfer in Continual Reinforcement Learning." *Advances in Neural Information Processing Systems* (NeurIPS), pp. 6304-6317, 2022.

[3] Wołczyk, Maciej, Karol J. Piczak, Bartosz Wójcik, Łukasz Pustelnik, Paweł Morawiecki, Jacek Tabor, Tomasz Trzcinski, Przemysław Spurek. "Continual Learning with Guarantees via Weight Interval Constraints." *International Conference on Machine Learning* (ICML), pp. 23897-23911, 2022.

[4] Wołczyk, Maciej, Bartosz Wójcik, Klaudia Bałazy, Igor T. Podolak, Jacek Tabor, Marek Śmieja, and Tomasz Trzcinski. "Zero Time Waste: Recycling Predictions in Early Exit ´ Neural Networks." *Advances in Neural Information Processing Systems* (NeurIPS), pp. 2516-2528, 2021.

[5] Wołczyk, Maciej, Magdalena Proszewska, Łukasz Maziarka, Maciej Zi¸eba, Patryk Wielopolski, Rafał Kurczab, and Marek Smieja. "PluGeN: Multi-Label Conditional ´ Generation From Pre-Trained Models." Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), pp. 8647-8656. 2022, CORE A*, 200 MEiN points.

Układ rozprawy

Rozprawa doktorska mgra Macieja Wołczyka składa się z dwóch części. W pierwszej części przedstawione są trzy publikacje, które dotyczą problemu uczenia ciągłego (ang. continual learning), ze szczególnym uwzględnieniem reżimu uczenia ze wzmocnieniem (ang. reinforcement learning). Te trzy publikacje w mojej stanowią trzon rozprawy, ze względu zarówno na ich bardzo wysoki poziom jak również spójność tematyczną.

W drugiej części rozprawy prezentowane są wyniki dotyczące adaptacji architektury sieci neuronowych w dwóch kontekstach: celem szybszego zakończenia obliczeń dla łatwych egzemplarzy wejściowych, a także dla modeli generatywnych aby umożliwić kontrolę nad charakterystyką generowanych próbek.

Uczenie ciągłe

Jednym z nierozwiązanych problemów uczenia maszynowego w ogólności jest zdolność adaptacji uzyskanych algorytmów do nowych danych i środowisk. W pracy [1] celem jest stworzenie ustandaryzowanego zbioru danych i środowisk w taki sposób, który umożliwi prowadzenie badań w dziedzinie uczenia ciągłego w sposób systematyczny, umożliwiający bezpośrednie porównanie wielu metod w tych samych warunkach. Wyniki ukazane w pracy [1] można podzielić na trzy kategorie:

- Przygotowanie środowiska ewaluacyjnego stanowiącego zbiór odpowiednio 10 i 20 zadań manipulacyjnych za pomocą robota Sawyer, w symulatorze Mujoco, środowisko to jest oparte na wcześniejszej implementacji Meta World. Zaproponowano dwa warianty: CW10 i CW20, gdzie w obu przypadkach używanych jest tych samych 10 zadań, jednakże w wariancie CW20 wszystkie zadania są powtórzone - motywacją jest tutaj możliwość sprawdzenia czy agent za drugim razem będzie w stanie szybciej opanować rozwiązanie zadania.
- Zaproponowanie nowego sposobu mierzenia szybkości uczenia nowych zadań, nazwany transferem w przód (ang. forward transfer). Zamiast porównywać jedynie wynik końcowy uzyskanego agenta, w zaproponowanej metodzie porównywana jest średnia skuteczność przez cały proces uczenia, dzięki temu w przypadku szybszego dojścia do takich samych wyników zaproponowana miara będzie w stanie pokazać różnicę pomiędzy podejściami.
- Ewaluacja istniejących metod w zaproponowanym środowisku. Zostały zewaluowane najbardziej popularne metody oparte o regularyzację, ponowne użycie danych z poprzednich zadań (ang. rehearsal) czy też wydzielenie pewnych części sieci neuronowej do konkretnych zadań, tak jak ma to miejsce w metodzie PackNet.

W mojej ocenie przedstawione środowisko może stać się standardem w pomiarze nowych wyników w dziedzinie uczenia ciągłego, co po części już widać w cytowaniach pracy, pomimo tego że została ona opublikowana mniej niż dwa lata temu.

W pracy [2] celem jest analiza i usprawnienie algorytmu SAC (Soft Actor Critic) tak aby uzyskać możliwie najlepsze wyniki dla uczenia ciągłego. Ewaluacja odbywa się na środowisku CW10/CW20 opracowanym w publikacji [1]. W algorytmie SAC można wyszczególnić kilka istotnych składników:

- aktor - model decydujący o podejmowanych akcjach,
- krytyk - model którego celem jest szacowanie wartości oczekiwanej nagród dla poszczególnych akcji,
- eksploracja - szczegóły dotyczące tego w jaki sposób agent zachęcany jest do odwiedzania przestrzeni stanów, które nie zostały jeszcze wystarczająco zbadane.

W pracy [2] przeanalizowane zostały wymienione składniki algorytmu SAC w różnych wariantach i ich wypływ na transfer podczas uczenia kolejnych zadań. W oparciu o tak wykonaną analizę zaproponowano algorytm CloneEx-SAC, który to algorytm przetestowano w środowiskach CW10/CW20 i wykazano jego wyższość nad algorytmami wcześniej testowanymi w pracy [1].

W pracy [3] celem jest opracowanie nowej architektury sieci neuronowej, która pozwala na uzyskiwanie matematycznych gwarancji utrzymywania skuteczności starych zadań nawet po przyuczeniu do nowych zadań w reżimie uczenia ciągłego. Główny pomysł pracy polega na zastosowaniu arytmetyki przedziałowej w przestrzenie parametrów sieci neuronowej, oraz utrzymywanie przestrzeni parametrów wewnątrz coraz węższych hiperprostokątów dla kolejnych trenowanych zadań. Dzięki temu, że hiperprostokąty są coraz węższe, uzyskane wcześniej gwarancje skuteczności nadal zachodzą, gdyż parametry mają jedynie coraz mniej swobody. Aby uzyskać taki efekt konieczne było przeprowadzenie arytmetyki przedziałowej we wszystkich komponentach sieci neuronowych. Warto zaznaczyć, że problem analizowany w pracy [3] jest bardzo trudny i z pewnością będzie wymagał dalszych badań, gdyż nawet wyniki przedstawione w pracy [3] mają swoje ograniczenia: nie stosują się do architektur zawierających pewne warstwy normalizujące takie jak batchnorm, a pomimo utrzymania gwaranancji na pesymistyczną skuteczność sieci, średnia skuteczność może maleć. Nie mniej jednak należy uznać, że przedstawione wyniki stanowią z pewnością krok we właściwym kierunku i umożliwiają dalszy rozwój badań.

Pozostałe wyniki

Tematem pracy [4] są modele, których czas działania jest różny w zależności od trudności przetwarzanego egzemplarza danych. Wykorzystywana jest w tym celu metoda wczesnego wyjścia (ang. early exit), gdzie po kolejnych warstwach sieci neuronowej dodatkowe moduły podejmują decyzję, czy wynik jest już na tyle pewny, że można zakończyć

przetwarzanie danych i uniknąć niepotrzebnych obliczeń, które byłyby wykonywane przez kolejne warstwy. Pomysł ten był już wykorzystywany we wcześniejszych pracach, jednakże artykuł [4] przedstawia dwa nowe pomysły:

- dodatkowe moduły podejmujące decyzję o wcześniejszym zakończeniu obliczeń nie są niezależne - dalsze moduły korzystają z wyników poprzednich,
- zwracany wynik szacujący pewność predykcji po kolejnych warstwach jest odpowiednio modyfikowaną średnią poprzednich wyników z wykorzystaniem trenowalnych parametrów.

Połączenie przedstawionych pomysłów z wcześniej stosowanymi architekturami pozwala na uzyskanie lepszych wyników, co zostało eksperymentalnie potwierdzone zarówno w dziedzinie klasyfikacji obrazów (CIFAR), ale również w reżimie uczenia ze wzmocnieniem dla algorytmu PPO w środowiskach Atari.

W pracy [5] przedstawiono moduł, który można połączyć z istniejącymi, już wytrenowanymi, modelami sieci generatywnych takich jak GAN czy VAE. Opracowany moduł ma za zadanie zmianę reprezentacji przestrzeni ukrytej w taki sposób, aby dało się sterować poszczególnymi cechami generowanych próbek (na przykład płeć, wiek, kolor włosów dla zdjęć postaci). Rzeczony moduł wymaga poetykietowania zbioru, który jest modelowany i etykiety te są wykorzystywane w procesie uczenia zmienionej reprezentacji za pomocą metody ang. *normalizing flow*.

Ocena rozprawy

W mojej ocenie przedstawione wyniki stanowią rozprawę doktorską na światowym poziomie - z dużym zapasem spełnia ona wymogi ustawowe i zwyczajowe stawiane rozprawom doktorskim. Badania dotyczą ważnych i aktualnych zagadnień, a wyniki zostały opublikowane na najlepszych światowych konferencjach z uczenia maszynowego i sztucznej inteligencji (wszystkie o rankingu A*). Co ważne, autor rozprawy jest wiodącym lub jednym z wiodących autorów w każdej z przedstawionych rozpraw.

Nie wyobrażam sobie aby rozprawa nie została wyróżniona, polecam również zgłoszenie rozprawy do Nagrody Prezesa Rady Ministrów za najlepsze rozprawy doktorskie.

Marek Cygan