

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Matyáš Lorenc
Název práce Evolution strategies for policy optimization in transformers
Rok odevzdání 2024
Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku Roman Neruda **Role** Vedoucí
Pracoviště Ústav informatiky AV ČR

Text posudku:

Ve své práci se student věnuje zkoumání optimalizačních algoritmů ze skupiny evolučních strategií v problému zpětnovazebného učení. Učený model je reprezentován moderní architekturou hluboké neuronové sítě Decision Transformer. Kombinace nových modelů hlubokých sítí a optimalizace používající konceptů novelty a quality diversity je originálním a aktuálním přístupem, který je navíc důležitý komplexní doménou zpětnovazebného učení.

Předkládaná práce je členěna do šesti kapitol včetně úvodu a závěru. Úvodní kapitola obsahuje motivaci, popis cílů a struktury dalšího textu. Ve druhé kapitole autor popisuje základy použitých technik, od popisu problému zpětnovazebného učení, přes evoluční strategie a koncepty novelty a quality diversity, až po transformery. Třetí kapitola je věnována aktuálnímu výzkumu a literatuře související s autorovým tématem. Vysvětlí se zde úspěšná metoda neuroevoluční strategie OpenAI-ES a její varianty pro práci s novelty a quality diversity NS-ES a NSR-ES, a v sekci 3.3 architektura Decision Transformer, specializovaného transformeru vhodného pro modelování politik agenta ve zpětnovazebném učení. Ve čtvrté kapitole autor popisuje svůj přístup využití popsáných variant evolučních strategií pro decision transformer včetně práce s předtrénovanými modely. Kapitola obsahuje i sekci s podrobnými informacemi o původní implementaci učících algoritmů. Rozsáhlá pátá kapitola popisuje řadu experimentů s variantami učení popsány výše. V šesté kapitole autor shrnuje dosažené výsledky a naznačuje možnosti další práce.

Za hlavní přínosy práce považují:

1. Originálním přínosem je základní přístup práce (popsaný v sekci 4.1) – využití evolučních strategií pro transformery v kontextu zpětnovazebného učení. Autor řeší spojení přístupů, které dosud v literatuře nebylo popsáno. Evoluční optimalizační techniky představují alternativu ke gradientním technikám a je velmi užitečné ověřit jejich efektivitu pro nové populární architektury hlubokých sítí. Některé z potenciálních výhod evolučních strategií jako robustnost nebo paralelizovatelnost jsou také ověřeny experimenty v kapitole 5.
2. Evoluční strategie jsou populačním optimalizačním algoritmem, který se snaží vyvážit exploraci a exploataci během prohledávacího procesu. Před několika lety byly navrženy varianty evoluce, které místo s účelovou funkcí pracují s koncepty originality nových řešení s překvapivě dobrými výsledky na problému učení chůze. Autor práce tyto koncepty novelty a quality diversity použil jako varianty klasické evoluční strategie a prozkoumal jejich úspěšnost.

3. Konečně, autor prozkoumal koncept předtrénování komplexního model decision transformeru pomocí dat získaných z jednodušších dopředných modelů (Algoritmus 4). Jako i předchozí varianty autorova přístupu, také tato je prakticky srovnána a ověřena v experimentech.

4. Experimentální část práce považuji za velmi rozsáhlou a pečlivě zpracovanou. Kromě popisu nastavení a výsledků experimentů obsahuje pátá kapitola také zevrubnou diskusi a komentáře.

Závěrem konstatuji, že předkládaná práce je velmi kvalitní, autor prokázal schopnost samostatné vědecké práce na aktuálním a originálním tématu strojového učení. Doporučuji dosažené výsledky k další publikaci. A rád doporučuji práci uznat jako diplomovou.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhujete na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum 5. února 2024

Podpis