

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Matyáš Lorenc

Název práce Evoluční strategie pro optimalizaci policy v transformerech

Rok odevzdání 2024

Studijní program Informatika - Umělá inteligence

Obor Informatika - Umělá inteligence

Autor posudku Mgr. Martin Pilát, Ph.D. **Role** oponent

Pracoviště KTIML MFF UK

Text posudku:

Cílem práce Matyáše Lorence bylo otestovat možnosti trénování neuronových sítí založených na transformerech pomocí evolučních strategií v oblasti zpětnovazebního učení. Tento cíl se podařilo splnit tím, že student otestoval několik typů evolučních strategií a několik způsobů učení neuronových sítí pomocí těchto strategií.

Práce je celkově rozdělena celkem do šesti kapitol (včetně úvodu a závěru). V prvních třech kapitolách napřed autor popisuje studovaný problém, metody, které se používají k jeho řešení, a související práce. Tato část práce je napsána dobře, poskytuje dobrý přehled existujících metod a obsahuje všechny informace potřebné pro zasazení práce do kontextu. V některých částech by mohla být napsána trochu podrobněji (např. v části o RNN chybí jakékoliv vzorečky popisující jejich fungování), autor ale odkazuje na vhodnou literaturu pro doplnění chybějících detailů a části, které jsou pro práci stěžejní, jsou popsány dostatečně detailně.

Čtvrtá kapitola popisuje postup, jakým autor dále zkoumá využití evolučních strategií pro trénování transformerů a popisuje i vytvořenou implementaci. Oceňuji zde podrobný rozbor rozdílů mezi originální implementací metod a tou vytvořenou autorem, včetně zdůvodnění veškerých rozdílů mezi těmito implementacemi.

Pátá kapitola obsahuje popis a výsledky všech provedených experimentů. Autor zde napřed vždy zkouší replikovat původní výsledky z původního OpenAI-ES článku a následně aplikuje metody pro transformery. Tento postup je velmi vhodný, neboť ukazuje, že i v případech negativních výsledků tyto jsou způsobeny architekturou sítě a ne jen samotným algoritmem. Výsledné porovnání je tak velmi přínosné a ukazuje jak silné, tak slabé stránky jednotlivých testovaných metod. Jen je škoda, že se z důvodu vysoké výpočetní náročnosti nepodařilo udělat delší experimenty, za kterých by mohly být vidět další zajímavé výsledky (jak i autor sám v práci zmiňuje).

Poslední kapitola potom shrnuje dosažené výsledky a poskytuje nápady na budoucí rozšíření.

Celkově je práce velmi dobře napsaná, obsahuje veškeré náležitosti, které by diplomová práce měla mít. Student ukázal, že je schopný pochopit a implementovat moderní metody hlubokého zpětnovazebního učení, experimentovat s nimi a experimenty patřičně vyhodnotit. Věřím, že výsledky práce jsou cenným přínosem do oblasti použití evolučních strategií pro trénování transformerů ve zpětnovazebním učení.

K práci mám pouze několik otázek k obhajobě:

Q1 V části 5.2.1 je navržena možná kombinace evolučních strategií a gradientních metod tak, že by se napřed použily evoluční strategie pro nalezení vhodné policy, která by se následně doladila gradientní (TD) metodou. Uvažovali jste i o možnosti použít TD metodu jako např. další mutaci v rámci evoluční strategie? Případně podobným způsobem prolínat evoluční a gradientní optimalizaci?

Q2 V části 5.3.3 se zmiňuje, že v počátečních fázích optimalizace je chování náhodného agenta preferované před chováním předtrénovaného agenta, protože má větší novost. Nedal by se tento problém jednoduše vyřešit tak, že by se do počáteční populace přidalo i několik náhodných agentů kromě agentů předtrénovaných?

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 5. února 2024

Podpis: