

Posudek na disertační práci Ing. Martina Bulína, M.Sc.

Multi-Agent based Neural Networks

**Doktorský studijní program, Studijní obor: Kybernetika
Západočeská univerzita v Plzni, Fakulta aplikovaných věd**

Pan Ing. Martin Bulín předložil disertační práci, která se zabývá učením neuronových sítí v multiagentním systému „Multi-Agent based Neural Networks“. Zvolené téma disertační práce patří do oboru disertace a je aktuální z hlediska současného stavu vědy.

Obsahem disertační práce je velmi důležitá oblast strojového učení, kterou jsou informační technologie. Konkrétně se jedná o využití umělých neuronových sítí při klasifikaci.

Práce je psána anglicky (úroveň jazyka nechci posuzovat, nejsem rodilá mluvčí), je členěna do 7 kapitol (včetně úvodu, závěru a 2 příloh) a je doplněna anotací resp. abstraktem (česky a anglicky), seznamem obrázků (98 položek), tabulek (8 položek), dále seznamem zkratk a symbolů, 138 položkami citované literatury z let 1980 až 2024, včetně prací nejvýznamnějších autorů zabývajících se umělými neuronovými sítěmi (McCulloch a Pitts, Rosenblatt, kritici perceptronu Minsky a Papert, Sejnowski, Hebb (1949), Hopfield, Kohonen, Werbos. Je také citováno 22 prací, jejichž je autorem nebo spoluautorem. Na základě přehledu vlastních publikací a publikací jejichž je spoluautorem uvedených v disertační práci mohu konstatovat, že podstatná část disertační práce byla publikována na odpovídající úrovni. Velké množství citací literárních pramenů považuji za velmi užitečné. Prostudované literární prameny nejsou zaměřeny jednostranně, ale jsou to převážně publikace v časopisech, dále knihy, články z Internetu, publikace na prestižních mezinárodních konferencích, abstrakty z článků, software, 3 technické zprávy a kvalifikační práce. Rozsah disertační práce je 133 stran.

V úvodu je uveden koncept řešení a jsou vyčteny cíle práce, tj. vysvětlení základních principů neuronových sítí z různých pohledů a možnosti navrhnout nové architektury pro řešení libovolných klasifikačních problémů.

V druhé nejrozsáhlejší kapitole jsou popsány základy z teorie neuronových sítí. Autor této doktorské práce uvádí stručný pohled do historie neuronových sítí. Velmi stručně se zmiňuje o biologických neuronových sítích, které byly inspirací pro zakladatele této oblasti výzkumu. Tato kapitola dále obsahuje přehled historických i současných typů neuronových sítí, jsou uvedeny principy architektur (včetně obrázků), u některých modelů i základní matematické vzorce. Uvedené architektury jsou dopředné a rekurentní sítě, transformery (nový přístup ke zpracování řad určený pro hluboké sítě používané pro aplikace zpracování textů) a speciální architektury. Jsou uvedeny základní učící algoritmy zaměřené spíše na sledování nedostatků zobecněných učících algoritmů, je popsán způsob hledání nových architektur, které se vztahují k předložené práci. Jsou popsány metody pro optimalizaci parametrů (počet vrstev, počet neuronů ve vrstvě, rychlost učení atd).

Ve třetí kapitole je popsán současný stav výzkumu v této oblasti. Autor se zmiňuje o principech a technologických postupech řešení především u podrobně rozebíraného posilovaného učení, na tomto místě práce se krátce zmiňuje o multiagentních systémech (zabývá se jimi podrobně ve 4. kapitole), velice stručně uvádí princip genetických algoritmů a human-in-the-loop systému, který je určen pro zlepšení účinnosti a přesnosti.

Čtvrtá kapitola by podle mého názoru měla být klíčová. Její název je zároveň názvem celé disertace. Podrobně je popsána mřížková struktura, je formulován problém a je zde uveden klíčový koncept a podrobně popsán algoritmus. Autor se rovněž zmiňuje o pěti možnostech rozšíření výzkumu v této oblasti.

V páté kapitole se autor věnuje vlastním experimentům (bylo realizováno pět klasifikačních problémů), jejich rozboru a vyhodnocení. Jako náročnou úlohu vidí možnost přizpůsobení úlohy libovolnému klasifikačnímu problému, což probírá v diskusi v šesté kapitole. Naznačuje budoucí směr výzkumu v této oblasti strojového učení, což demonstruje na osmi námětech.

Ve velmi stručném závěru v sedmé kapitole autor disertace shrnuje cíl svého výzkumu, kterým je návrh kompaktní architektury neuronové sítě a hledání významu jejich individuálních komponent.

Po formální stránce je předkládaná práce na dobré úrovni, je psána přehledně. Menší kvalitu mají některé obrázky. Jsou částečně neostře (nejspíš zkopírované z jiných publikací) doplněné ostrým vlastním popisem. U obrázků nejsou většinou uvedeny citace. Znamená to, že jsou vytvořeny autorem? K úrovni vlastního textu je třeba konstatovat, že metodologie je nová a neobjevila se v publikacích jiných autorů. Obsahuje několik velmi zajímavých konceptů a strategií strojového učení. Vlastní přínos vidím v alternativním pohledu na využití neuronových sítí. Pan inženýr Bulín navrhl vlastní metodu pro návrh architektury umělých neuronových sítí pro klasifikační úlohu. Stav výzkumu sice není úplně popsán, ale výzkum ještě není ukončen. Je možné předpokládat, že v blízké budoucnosti budou jeho výsledky hrát velmi významnou roli v rozvoji klasifikačních metod.

K práci mám následující připomínky a dotazy:

Připomínky:

- Str. 51, kap. 2.7: nesouhlasím s vyjádřením, že v předchozích sekcích (do 2.6) byl podán „úplný“ přehled ANN architektur. Podle mého názoru jste se zaměřil především na dopředné a rekurentní sítě. O velmi významných SOMech (zejména pro některé oblasti aplikací) a o jejich význačných vlastnostech jste se zmínil jen v několika řádcích.

Dotazy k práci:

- Jsou možné i jiné typy rekurencí u rekurentních sítí?
- K čemu jsou vhodné Elmanovy rekurentní sítě a jak v nich mohou být vedeny rekurentní vazby?
- Jaké jsou výhody aplikací umělých neuronových sítí a jaké jsou jejich nevýhody?

Na závěr konstatuji, že vytčené cíle byly splněny, předložená disertační práce je zajímavá, odpovídá kvalitě, která je požadovaná pro tento typ práce. Za získanými výsledky se skrývá velmi mnoho užitečné práce.

Předložená disertační práce odpovídá, podle mého soudu, obecně uznávaným požadavkům na udělení akademického titulu Ph.D. Pan Ing. Bulín projevilschopnost podílet se na výzkumu, z předložené disertační práce vyplývá jeho schopnost samostatně vědecky pracovat.

Disertační práci doporučuji k obhajobě.

V Praze 21. 5. 2024

Prof. Ing. Jana Tučková, CSc.

Posudek disertační práce Ing. Martina Bulína, MSc.

Na téma „Multi-Agent based Neural Networks“

Práce pana Bulína se zabývá možností kombinovat metody umělých neuronových sítí, multiagentních systémů, genetických algoritmů a zpětnovazebního učení (Reinforcement Learning). Kombinovaná metoda je určena k návrhu topologie řídké neuronové sítě, schopné řešit zadanou úlohu klasifikace s minimálním množstvím synapsí.

Teoretická část práce (kapitoly 2 a 3) shrnuje dosavadní relevantní znalosti o umělých neuronových sítích, multiagentních systémech, zpětnovazebním učení a genetických algoritmech. Většina této části je velmi kvalitně vypracovaná, podrobná a dobře podložená citacemi. Především některé části o neuronových sítích by byly dobrým základem samostatné učebnice posledního vývoje v oblasti NS. Drobným nedostatkem je kapitola 3.3 (genetické algoritmy), která je poměrně povrchní a postrádá citace.

Hlavní část práce pak popisuje vlastní navržený systém, jeho fungování a experimentální výsledky. Autor úspěšně zkombinoval všechny uvažované postupy, tedy neuronové sítě, multiagentní systémy zpětnovazebního učení a genetický algoritmus. Vytvořený systém pracuje s potenciálně plně propojenou, dopřednou neuronovou sítí, kde každá synapse je kontrolována agentem. Agent může v rámci návrhu topologie synapsi ze sítě vyjmout, nebo ji do sítě vložit. Chování agenta je trénováno zpětnovazebním učením. Nejlépe natrénovatelné topologie neuronové sítě jsou vybírány genetickým algoritmem.

Systém byl následně otestován na pěti úlohách různé složitosti s velmi nadějnými výsledky. Ve všech případech našel systém velmi řídké sítě, které přesto dosahovaly vysoké (většinou 100%) přesnosti. Tato část práce bohužel neobsahuje ani orientační informace o výpočetní náročnosti řešení, i když autor rovnou připouští, že časová náročnost existující implementace je hlavní překážkou nasazení systému na rozsáhlejší úlohy. Autor upozorňuje, že efektivita implementace nebyla prioritou předložené práce a vyjadřuje naději, že je možné implementaci výrazně zrychlit.

Práce obsahuje seznam řady vlastních publikací a spolupublikací autora v oblasti umělé inteligence, hlavně neuronových sítí a zpracování přirozené řeči a jazyka. Zřejmý je dlouhodobý zájem autora o řídké neuronové sítě. Seznam dokládá zkušenost autora s širokou řadou úloh umělé inteligence a jeho schopnost prezentovat výsledky práce na předních domácích i zahraničních konferencích.

Z formálního hlediska je práce kvalitně napsaná, systematická a srozumitelná bez podstatných chyb. Všechny podstatné části textu jsou podloženy citacemi z rozsáhlého seznamu použité literatury. Porozumění textu částečně komplikuje zvyk autora nechávat si nejdůležitější informace na konec odstavce, kapitoly nebo celého textu.

Autor se pochopitelně nevyhnul drobným formálním nedostatkům jako například:

1. Rovnice 3.8 a několik následujících obsahují pravděpodobně typografické chyby, protože říkají, že $a = a + b$ kde $b \neq 0$.

2. Text obsahuje několik vět, které se mi nepodařilo pochopit (například zmínka o „floppy disc“ na straně 92.
3. Není popsáno, co představuje zelená křivka v grafech 5.3 a 5.11

K navržené metodě mám dále několik dotazů:

1. Při použití genetických algoritmů autor používá pouze operaci řízené mutace bez náhodné mutace a křížení. Je takto omezený přístup použitý i jinde v literatuře nebo se jedná o modifikaci navrženou jen pro tuto práci?
2. V sekci 4.2.3 je použita analogie s multiagentním systémem, avšak agenti v předloženém řešení spolu interagují jen formou vzájemně závislých vstupních vektorů. Je tento přístup používán i jinde nebo se jedná o unikátní zjednodušení?
3. Na obrázku 4.17 je uvedeno, že jádro agentů a kritika je implementováno vícevrstevným perceptronem. Jaké jsou rozměry této sítě a jak byly tyto parametry zvoleny?
4. Bylo by možné upřesnit, jak je definován vstupní parametr *apriory*? Z obrázku 4.19 to není zcela zřejmé.
5. Monitor popsáný v kapitole 4.3.3 je součástí práce autora nebo se jedná o část systému PyTorch zmíněný v kapitole 2.8?
6. V kapitole 5.1 je uvedeno, že jedním z výstupů algoritmu je natrénovaný agent pro vyjímání a vkládání synapsí. Existuje nějaká souvislost mezi chováním agentů natrénovaných na různých úlohách nebo je toto chování specifické pro každou úlohu?

Předložená práce je velmi zajímavým, původním a inovativním příspěvkem v oboru neuronových sítí. Autor si vytýčil velmi ambiciózní cíl – předložit alternativu ke všeobecně používaným, masivním, plně propojeným sítím s miliony parametrů – a tento cíl splnil. Při tom prokázal dobrou znalost použitých technologií a schopnost je efektivně využívat a inovativně kombinovat, modifikovat a zjednodušovat. Proto doporučuji práci k obhajobě.

26. 5. 2024

Ing. Petr Bečvář Ph.D.