



УНИВЕРЗИТЕТ У БЕОГРАДУ - ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ

Булевар краља Александра 73, 11000 Београд, Србија

Тел. 011/324-8464, Факс: 011/324-8681

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 04.09.2018. године именовало нас је у Комисију за преглед и оцену мастер рада дипл. инж. Богдане Нонковић под насловом „Играње аркадних игара применом дубоког учења са подстицањем“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Богдана Нонковић је рођена 23.06.1994. године у Книну, Хрватска. Осмогодишње школовање је започела у основној школи „Никола Тесла“, наставила у школи „Стеван Синђелић“ а завршила у основној школи при „Математичкој гиманзији“. Даље школовање је наставила у „Математичкој гиманзији“ коју је завршила са одличним успехом. Електротехнички факултет у Београду је уписала 2013. године. Дипломирала је на одсеку за Сигнале и системе 2017. године са просечном оценом 9,49. Дипломски рад на тему „Имплементација Ехрестимах алгорита на примеру Расман игрице“ одбранила је у септембру 2017. године. Даље школовање је наставила на Електротехничком факултету где је на академским-мастер студијама положила испите са просечном оценом 9,8. Током периода образовања активно се бавила шахом, остварујући запажене резултате како у појединачном тако и у екипном такмичењу.

2. Опис мастер рада

Мастер рад обухвата 43 стране, са укупно 13 слика, 2 табеле и 18 референци. Рад садржи 6 поглавља (укључујући закључак), прилог са написаним кодом и списак коришћене литературе.

Прво поглавље представља увод у коме су описани предмет и циљ рада: имплементација алгорита који омогућава софтверском агенту да самостално научи да игра видео игру, користећи само сирове пикселе са конзоле и тренутни резултат као улазе, без експлицитног програмирања правила игре нити стратегија које агент треба да користи. Дат је кратак преглед основне идеје која је примењена у раду и окружења које је коришћено за имплементацију и тестирање алгорита.

У другом поглављу су описани теоријски основи алгорита. Дефинисан је проблем учења подстицањем на супрот стандардном машинском учењу са надзором и учењу без надзора. Описан је Марковљев процес одлучивања (МПО) као основни математички модел у проблемима учења подстицањем. Затим је детаљно објашњен појам Q-учења на основу временских разлика као приступ који се може користити у случајевима када модел окружења није унапред познат. У наставку је описана конволуциона неурална мрежа, стандардни алат машинског учења у случајевима када је улаз представљен сликом. Коначно су детаљно описани архитектура, хипер-параметри и начин обучавања дубоке Q-неуралне мреже, која је у овом раду коришћена као нелинеарни апроксиматор Q-вредности појединачних стања.

Треће поглавље садржи опис видео игре која је коришћена у сврху валидације имплементације алгорита. Ради се о игри Понг, поједностављеној симулацији стоног тениса. Коришћено је окружење из библиотеке OpenAI Gym.

У четвртном поглављу су детаљно изложени детаљи имплементације. Најпре је објашњено на који начин је слика процесирана како би се учинила погоднијом за сврху

учења подстицањем. Објашњени су појмови понављања искустава и увођења засебне, циљне неуралне мреже. Ове две технике су кључне за стабилизацију процеса учења. Анализирани су меморијски захтеви алгорита у погледу параметара саме мреже као и у погледу величине бафера у ком се чувају претходна искуства.

Пето поглавље садржи опис остварених резултата. Приказани су графици који сведоче о томе да учење конвергира као неком стационарном закону управљања. Анализирани су поступци агента у неким карактеристичним ситуацијама током игре и указано је на смисленост одабраних акција.

Последње поглавље садржи резиме остварених резултата и преглед најзначајнијих резултата из литературе који се могу применити у циљу побољшања имплементираних алгорита кроз будући рад.

3. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Богдане Нонковић се бави проблемом дубоког учења са подстицањем. Конкретно, у раду је описано и имплементирано тзв. дубоко Q-учење, где се користи конволуциона неурална мрежа као апроксиматор Q-вредности. Алгоритам је верификован применом на видео игри Понг из Atari окружења библиотеке OpenAI Gym.

Кандидаткиња је темељно прегледала релевантну савремену стручну литературу. Полазна тачка је данас познати научни рад истраживача из фирме DeepMind из 2013. године (референциран у тези), где је први пут описан алгоритам дубоког Q-учења. После ове публикације дошло је до интензивирања истраживања у области дубоког учења са подстицањем и објављивања великог броја радова, од којих су најзначајнији укратко описани и референцирани у тези. Кандидаткиња је имплементирала основни алгоритам уз нека побољшања предложена у каснијим публикацијама и темељно га верификовала кроз симулације. Предности и недостаци алгорита су критички анализирани, и указано је на могуће правце даљег истраживања.

Основни доприноси рада су: 1) детаљан преглед и критичка анализа литературе из области дубоког учења са подстицањем, са акцентом на дубоком Q-учењу; 2) имплементација алгорита дубоког Q-учења и његова верификација на једном од стандардних проблема који се помињу у литератури; 3) дате су перспективне идеје за могући даљи рад на овој теми.

4. Закључак и предлог

Кандидаткиња Богдана Нонковић је у свом мастер раду успешно применила алгоритам дубоког Q-учења у циљу имплементације агента који самостално учи да игра видео игре. При томе је исказала самосталност, систематичност и иновативност у своме раду, као и способност коришћења релевантне литературе.

На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Богдане Нонковић прихвати као мастер рад и кандидаткињи одобри јавну усмену одбрану.

Београд, 10.09.2018. године

Чланови комисије:

П. Тадић

др Предраг Тадић, доцент

Горан Квашчев
др Горан Квашчев, ванредни професор