



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 09.09.2019. године именовало нас је у Комисију за преглед и оцену мастер рада дипл. инж. Душана Попадића под насловом „Управљање инверзним клатном путем DDPG и А3С алгоритама учења подстицањем“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Душан Попадић је рођен 29.12.1995. године у Београду. Рачунарску гимназију је завршио у Београду као носилац дипломе „Вук Караџић“ и ученик генерације. Електротехнички факултет у Београду уписао је 2014. године. Дипломирао је на модулу Сигнали и системи 2018. године са просечном оценом 9,59. Дипломски рад је одбранио у септембру 2018. године са оценом 10. Мастер академске студије на Електротехничком факултету у Београду је уписао 2018. на модулу Сигнали и системи. Положио је све испите са просечном оценом 10.

2. Опис мастер рада

Мастер рад обухвата 38 страна, са укупно 22 слике, 1 табелом и 11 референци. Рад садржи увод, 5 поглавља и закључак (укупно 7 поглавља) и списак коришћене литературе.

У уводу су дати предмет, циљ и очекивања од мастер рада. Укратко је описана идеја и могућности примене учења подстицањем, а дат је и преглед литературе.

У другом поглављу је дат теоријски увод у учење подстицањем, објашњени су Марковљеви процеси одлучивања и описани су главни приступи у учењу подстицањем. Посебно је објашњен актер-критичар приступ, као и приступ заснован на градијенту управљања и описана је теорема о градијенту управљања.

У трећем поглављу је описано симулационо окружење са којим ће интераговати имплементирани алгоритми, његова стања, могуће акције и начин доделе награде.

Четврто поглавље детаљно описује DDPG алгоритам, његову математичку позадину и имплементацију. Дате су конкретне структуре актер и критичар модела имплементираног алгоритма као и изабрани параметри.

Пето поглавље детаљно описује А3С алгоритам, његову математичку позадину и имплементацију. Дате су конкретне структуре актер и критичар модела имплементираног алгоритма као и изабрани параметри.

У шестом поглављу су представљени, анализирани и упоређени резултати симулација.

У седмом поглављу су сумирани закључци настали на основу добијених резултата, поређење добијених резултата са очекивањима која су наведена у уводу и предлог за надоградњу рада.

3. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Душана Попадића се бави анализом два популарна алгоритма учења подстицањем: *Deep Deterministic Policy Gradient (DDPG)* и *Asynchronous Actor-Critic Agents (A3C)*. Учење подстицањем је подобласт машинског учења у којој алгоритам (у

литератури је уобичајен назив „агент“) треба да научи оптималну стратегију за обављање неког задатка кроз директну интеракцију са окружењем. У овом раду, окружење је симулирано усправно клатно монтирано на колица која могу да се крећу праволинијски, а задатак је да се управља брзином колица тако да клатно стоји усправно. DDPG и A3C су објављени 2016. године али су брзо постали веома популарни у научној заједници и данас важе за референтне методе у овој области.

Кандидат је детаљно проучио релевантну литературу и дао прецизан опис ових алгоритама у свом раду, заједно са детаљима сопствене имплементације у програмском језику Python. Спроведене симулације су темељно анализирани а алгоритми су међусобно упоређени према брзини конвергенције и квалитету научене политике (закон управљања који пресликава стања у акције), као и према нумеричкој комплексности, могућности паралелизације и проблемима при практичној имплементацији, имајући у виду ситуације у којима би агенти интераговали са реалним, физичким окружењем, за разлику од симулираног што је овде био случај.

Основни доприноси рада су: 1) приказ и методологија имплементације DDPG и A3C алгоритама учења подстицањем; 2) примена имплементираних алгоритама на симулираном инверзном клатну уз темељно поређење перформанси двају алгоритама по више критеријума; 3) уочене могућности за побољшања и наставак рада у области.

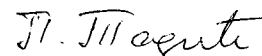
4. Закључак и предлог

Кандидат Душан Попадић је у свом мастер раду успешно применио два стандардна алгорита учења подстицањем на проблем управљања симулираним инверзним клатном, критички анализирао њихове перформансе и уочио могућности за даље истраживање. При томе је исказао самосталност и систематичност у раду и демонстрирао способност коришћења савремене литературе из ове области.

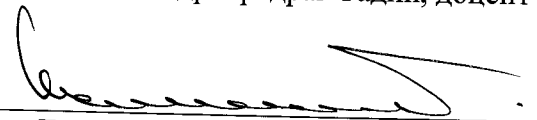
На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Душана Попадића прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 12. 09. 2018. године

Чланови комисије:



др Предраг Тадић, доцент



др Горан Квашчев, ванредни професор