

**МОЛБА
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА**

Молим да ми се одобри израда мастер рада под насловом:

Учење поткрепљивањем са дистрибуираном архитектуром

Значај теме и области:

Проблеми које учење поткрепљивањем (енг. *reinforcement learning*) решава су уобичајено представљени као учење понашања агента који врши интеракцију са неким динамичким окружењем чије се акције награђују или кажњавају. Циљ алгоритма је учење стратегије агента којом се доноси секвенца одлука тако да максимизују дугорочну награду. Поред бројних изазова која ова парадигма машинског учења пружа, веома чест је и проблем недовољног броја улазних података или њиховог неефикасног генерисања. Идеја дистрибуираних алгоритама учења поткрепљивањем је да се део за скупљање/генерисање искустава и део за учење модела потпуно раздвоји. Примена дистрибуиране архитектуре ствара могућност паралелног извршања великог броја агената у својим засебним окружењима чија се искуства могу искористити за учење једне централизоване мреже. Паралелизовањем агената се добија широк спектар могућности за модификовање и проширивање већ постојећих алгоритама. Могу се истовремено користити различите додатне технике при интеракцији агената са окружењем, чиме би се потенцијално повећала експлоративност учења модела. Могу се скупљати искуства свих агената у једном баферу где ће се на основу предефинисаних метрика додељивати приоритет или оцена подацима искуства. Број побољшања које увођење дистрибуираности омогућује је веома велик и постоји велики број техника које су знатно унапредиле добијене резултате у учењу поткрепљивањем. Значајан део најмодернијих алгоритама прати дистрибуиране архитектуре.

Специфични циљ рада:

Циљ рада јесте преглед дистрибуиране архитектуре код алгоритама дубоког учења поткрепљивањем. Анализираће се улога, скалабилност, изазови и решења појединачних компоненти такве архитектуре. Имплементација ће бити заснована на архитектури тока догађаја (енг. *event-streaming*) и биће примењена на решавање проблема играња видео игара. Постигнути резултати у самим видео играма, као и сама ефикасност и брзина учења, ће бити упоређени са резултатима постигнутим са недистрибуираним моделима.

Никола Ковачевић, 1104/2019, Информатика
(име и презиме студ., бр. инд., ознака програма и модула)

(својеручни потпис студента)

(датум подношења молбе)

Катедра _____ је сагласна са предложеном темом.

(шеф катедре)

Сагласан ментор

проф. др Владимир Филиповић

(својеручни потпис ментора)

Чланови комисије

1. доц. др Младен Николић

2. доц. др Александар Картељ

(датум одобравања молбе)