

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Милоша Котлара

Одлуком 5003/17-3 бр. од 04.03.2022. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Милоша Котлара под насловом

Детекција аномалија коришћењем мета података у аутоматизованим системима за машинско учење

После прегледа достављене Дисертације и других пратећих материјала и разговора са Кандидатом, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Милош Котлар је уписао докторске академске студије на Електротехничком факултету Универзитета у Београду, модул Рачунарска техника и информатика, школске 2017/2018 године. На докторским студијама положио је све испите са просечном оценом 10,00.

02.09.2021. године кандидат је пријавио тему за израду докторске дисертације под радним називом „Детекција аномалија коришћењем мета података у аутоматизованим системима за машинско учење”.

07.09.2021. године Комисија за студије трећег степена разматрала је предлог теме за израду докторске дисертације и предлог Комисије о оцени научне заснованости теме упутила Наставно–научном већу на усвајање.

Наставно-научно веће именовало је Комисију за оцену научне заснованости теме (Одлука бр. 5003/17-1 од 24.09.2021. године) у саставу:

- др Милош Цветановић, ванредни професор, Електротехнички факултет у Београду
- др Горан Квашчев, ванредни професор, Електротехнички факултет у Београду
- др Сениша Влајић, редовни професор, Факултет организационих наука у Београду

За менторе докторске дисертације су предложени др Марија Пунт, доцент и др Захарије Радивојевић, ванредни професор.

08.10.2021. године обављена је јавна усмена одбрана предложене теме докторске дисертације на Електротехничком факултету у Београду пред наведеном Комисијом. Комисија је закључила да је кандидат добио оцену „задовољио”.

Наставно-научно веће усвојило је Извештај Комисије за оцену научне заснованости теме докторске дисертације (Одлука бр. 5003/17-2 од 16.11.2021. године).

Веће научних области техничких наука дало је сагласност на предлог теме докторске дисертације (број 61206-4938/2-21 од 10.12.2021. године).

10.02.2022. године кандидат је предао докторску дисертацију на преглед и оцену под насловом „Детекција аномалија коришћењем мета података у аутоматизованим системима за машинско учење”.

14.02.2022. године Комисија за студије трећег степена потврдила је испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације.

Наставно-научно веће Електротехничког факултета именовало је Комисију за преглед и оцену докторске дисертације (Одлука бр. 5003/17-3 бр. од 04.03.2022. године) у саставу:

- др Милош Цветановић, ванредни професор, Електротехнички факултет у Београду
- др Драган Бојић, редовни професор, Електротехнички факултет у Београду
- др Сениша Влајић, редовни професор, Факултет организационих наука у Београду

која је на истој седници проширена чланом:

- др Горан Квашчев, ванредни професор, Електротехнички факултет у Београду

1.2. Научна област дисертације

Докторска дисертација „Детекција аномалија коришћењем мета података у аутоматизованим системима за машинско учење” кандидата Милоша Котлара припада научној области Електротехника и рачунарство, ужа научна област Рачунарска техника и информатика, за коју је матичан Електротехнички факултет (ЕТФ) - Универзитет у Београду.

Ментори докторске дисертације су др Марија Пунт, доцент Електротехничког факултета у Београду и др Захарије Радивојевић, ванредни професор Електротехничког факултета у Београду.

1.3. Биографски подаци о кандидату

Милош Котлар је рођен 07.11.1993. године у Београду. Електротехнички факултет уписао је 2012. године и дипломирао на модулу софтверско инжењерство 2016. године са просечном оценом 9,24, одбравивши дипломски рад код проф. др Вељка Милутиновића. Дипломске академске – мастер студије на Електротехничком факултету у Београду уписао је 2016. године и мастерирао на модулу софтверско инжењерство 2017. године са просечном оценом 9,50, одбравивши мастер рад код доц. др Марије Пунт. Докторске студије уписао је 2017. године на модулу рачунарска техника и информатика и положио је све испите са просечном оценом 10,00. Од 2016. године био је запослен као софтверски инжењер у фирми „АВВ”, Цирих, Швајцарска. Од 2019. године запослен је као водећи истраживач у фирми

„TraceLabs”, Љубљана, Словенија.

Милош Котлар је научно-истраживачко искуство започео 2016. године у сарадњи са Електротехничким факултетом у Београду на радовима у области *dataflow* рачунара и машинског учења. Даље искуство је стекао учествовањем на два пројекта као водећи истраживач у фирми „TraceLabs”, финансираним од стране ЕУ у оквиру *H2020* програма. Пројекти у којима је учествовао припадају области машинског учења, проналажења скривеног знања, графова знања, дистрибуираних система, и архитектуре рачунара.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација је написана на српском језику, ћириличним писмом, садржи 126 страна (99 страна текста), 36 слика, 30 табела и 79 библиографских референци. Састоји се од насловне стране на српском и енглеском језику, стране са информацијама о менторима и члановима комисије, захвалнице, кратког резимеа на српском и енглеском језику и 6 поглавља:

1. Увод (2 стране)
2. Дефиниција проблема и преглед постојећих решења (29 стране)
3. Предлог решења (9 стране)
4. Пројектовање компоненте за одабир алгорита у аутоматизованим системима за машинско учење (14 стране)
5. Евалуација предложеног решења и имплементираних компоненти (43 стране)
6. Закључак (2 стране)

Такође, докторска дисертација садржи списак коришћене литературе, листу скраћеница, листу слика, листу табела, биографију аутора, изјаву о ауторству, изјаву о истоветности штампане и електронске верзије докторског рада и изјаву о коришћењу.

2.2. Кратак приказ појединачних поглавља

У **првом поглављу** уводе се основне информације о истраживачкој области и појмови неопходни за дефинисање проблема приликом одабира алгорита за детекцију аномалија. Дати су уопштено циљ и допринос рада, представљени су добијени резултати и дат је преглед наредних поглавља.

У **другом поглављу** дефинисан је домен истраживања, сви потребни појмови за разумевање проблема и извршена је упоредна анализа постојећих решења у домену мета учења и мета података за одабир алгорита за детекцију аномалија у подацима. Поглавље даје преглед типова и домена података у којима се јављају аномалије, преглед различитих типова аномалија, алгорита за детекцију аномалија и аутоматизованих система за машинско учење. Након тога је дат преглед и класификација постојећих решења у домену мета учења кроз различите аспекте комплексности. На крају су анализиране функције које се користе за мерење сличности између мета података коришћењем функција за мерење удаљености.

У **трећем поглављу** дефинисан је скуп критичних захтева које функције за израчунавање мета података морају да задовоље како би могле да се користе у аутоматизованим системима за детекцију аномалија. Након тога, дефинисане су полазне хипотезе на које овај рад треба да одговори. Затим је дат предлог функција за израчунавање мета података заснованих на доменском знању које могу да се користе у аутоматизованим системима за детекцију аномалија.

У четвртом поглављу дефинисана је логичка структура компоненте за одабир алгорита у системима за аутоматизовано машинско учење, која се заснива на предложеном скупу функција за израчунавање мета података. Након тога је описан начин функционисања такве компоненте и представљена је топологија компоненте у различитим окружењима.

У петом поглављу представљена је поставка експеримената, укључујући скупове података, евалуационе метрике, алгоритме за детекцију аномалија, функције за мерење сличности између мета података и дат је преглед коришћених архитектура за извршавање експеримената у различитим окружењима. Након тога је дат опис експеримената који проверавају дефинисане полазне хипотезе. Добијени експериментални резултати су анализирани на начин да се одговори на постављене полазне хипотезе.

У шестом поглављу су сумирани резултати приказани у докторској дисертацији, изложен је закључак и могући правци даљег истраживања.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Детекција аномалија представља битан фактор приликом анализе података, директно утиче на квалитет података и има велику примену у различитим областима. Најчешћи случајеви детекције аномалија у подацима су за потребе праћења квалитета података, и заступљени су у доменима транспорта, медицине, софтверских записа, интернет безбедности и другим. Аномалије могу да се појаве у различитим типовима података и потребно их је детектовати као грешке у подацима или као нову категорију. У зависности од типа и домена података у ком се детектују, аномалије могу да имају различите карактеристике и могу да представљају грешке или нове категорије у подацима, што чини њихову детекцију могућом коришћењем одговарајућег модела.

Потреба за детекцијом аномалија захтева креирање модела који ће моћи ефикасно да детектује аномалије у подацима. Такав модел може да се разликује у зависности од домена примене, при чему одабир модела и параметара модела за детекцију аномалија искључиво зависи од експертизе креатора система или доменског експерта. Процес одабира модела и параметара модела може да се аутоматизује коришћењем аутоматизованих система за машинско учење (*AutoML*). Такви системи решавају проблем одабира модела и омогућавају кориснику који није експерт из одређеног домена, или нема довољно доменског знања, да креира модел за детекцију аномалија без претходног знања о карактеристикама, имплементацији и начину рада модела. Како би аутоматизовани систем за машинско учење могао да одабере модел који даје задовољавајуће резултате приликом детекције аномалија за дате податке и оптимизациону метрику, неопходно је применити савремене методе мета учења за одабир одговарајућег модела.

Мета учење представља савремену методу машинског учења где се применом одговарајућих функција над подацима добијају мета подаци који су предиктивни за одабир модела заснованог на алгоритмима машинског учења. Мета подаци се израчунавају применом функција над скуповима података, и на тај начин представљају њихове карактеристике. У аутоматизованим системима за машинско учење, мета подаци се користе приликом одабира модела на основу претходно стеченог знања и података који систем евалуира. Како би се креирале функције за израчунавање мета података, неопходно је одредити карактеристике аутоматизованих система за машинско учење и на основу тога дефинисати критичне захтеве које функције морају да испуњавају. На тај начин је могуће дефинисати скуп функција за израчунавање мета података које испуњавају критичне захтеве потребне за коришћење у аутоматизованим система за детекцију аномалија. Критични захтеви су условљени окружењем у ком се систем налази као и архитектуром на којој се

систем извршава, и могу да се односе на стриктна временска ограничења као и на доступне ресурсе којима систем располаже.

Оригиналност овог рада у домену мета учења представља нови приступ за израчунавање мета података на основу доменског знања, где мета подаци описују карактеристике аномалија у подацима. Предложени скуп мета података заснован на доменском знању за детекцију аномалија није разматран у постојећој литератури других аутора и представља оригиналан допринос објављен у радовима кандидата. Предложено решење има значајно мању комплексност у односу на постојећа решења, испуњава критичне захтеве за примену у аутоматизованим системима за машинско учење и могуће је проценити мета податке искључиво на основу доменског знања. У домену аутоматизованих система за машинско учење, предложено решење се фокусира на побољшање компоненте за одабир алгоритма, које решава проблем са перформансама и скалабилношћу система, као и коришћења на различитим локацијама извршавања.

3.2. Осврт на референтну и коришћену литературу

У оквиру дисертације коришћена је литература која броји 79 библиографских јединица из више области. Референце обухватају области којима се бави дисертација:

- аномалије у подацима
- алгоритми за детекцију аномалија
- аутоматизовани системи за машинско учење
- локације извршавања и архитектуре
- мета учење
- алгоритми за мерење удаљености

Може се сматрати да дисертација садржи детаљан и актуелан преглед адекватне литературе. У случају одређених референци дате су и појединости које реферисани радови доносе и упоређене су са карактеристикама решења предложеног у овој дисертацији.

Списак литературе коју је кандидат навео показује да је извршио детаљну и свеобухватну анализу из више области које се односе на предмет рада. Наведена литература обухвата разне врсте научних резултата попут оригиналних истраживања радова из водећих часописа и конференција, прегледне научне радове и изворе са Интернет страница. У склопу цитиране литературе су и ауторски радови кандидата на којима се заснивају и неки од резултата представљених у дисертацији.

3.3. Опис и адекватност примењених научних метода

Током истраживања у оквиру ове докторске дисертације, примењене су следеће научне методе:

- Анализа постојећих скупова података који садрже аномалије у подацима
- Анализа аномалија по различитим карактеристикама од интереса, као што су тип аномалија, локалитет аномалија и димензионални простор аномалија
- Анализа различитих алгоритама за детекцију аномалија
- Преглед различитих аутоматизованих система за машинско учење
- Преглед и анализа постојећих решења мета учења за одабир алгоритма
- Имплементација компоненте за одабир алгоритма коришћењем мета података
- Анализа утицаја локације извршавања на одабир архитектуре за имплементацију компоненте за одабир алгоритма у домену детекције аномалија
- Евалуација перформанси предложеног решења и имплементираних компоненти кроз аспекте од интереса

Научне методе примењене у дисертацији су заступљене у теоријским и експерименталним истраживањима у области машинског учења и аутоматизованих система. На основу свега наведеног може се закључити да спроведене научне методе на адекватан начин одговарају циљевима који су постављени на почетку истраживања.

3.4. Применљивост остварених резултата

Предложено решење и добијени резултати у дисертацији имају практичну примену за одабир алгорита у аутоматизованим системима за детекцију аномалија. Имплементирана компонента за одабир алгоритама представља целину у аутоматизованом систему за детекцију аномалија коју је могуће интегрисати у постојеће или нове система за детекцију аномалија. Приступ коришћен у раду за одређивање мета података на основу доменског знања може да се примени и на друге области како би се проверило да ли се доменским мета подацима могу карактерисати подаци тако да се врши предлагање алгоритама у аутоматизованим системима за машинско учење. Упоредна анализа алгоритама који се користе за детекцију аномалија, креирани репозиторијум података, предложене функција за мерење удаљености и предложени мета подаци представљају добру основу за даља истраживања у овој области.

Даља истраживања у овој области укључују проширивање скупа мета података за детекцију новим доменским и простим мета подацима, као и тражењем минималног скупа комбинацијом постојећих решења и предложеног решења. Такође, увођењем мета података заснованих на структурним и вредносним ограничењима скупова података може додатно да смањи комплексност израчунавања мета података и представља добар правац за будућа истраживања у овој области.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Ова дисертација је направила искорак у примени аутоматизованих система за машинско учење у детекцији аномалија. Оваква примена укључује широк сегмент технологија, укључујући алгоритме машинског учења и мета учење. Спроведено истраживање захтевало је теоријски увид у наведене области, практичну реализацију и проверу постављених хипотеза кроз дефинисање, имплементирање, тестирање и евалуацију предложеног решења. На основу тога, можемо закључити да је кандидат Милош Котлар показао задовољавајући степен способности за самостални научни рад и реализацију практичних инжењерских решења.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Прототип аутоматизованог систем за детекцију аномалија је пројектован и имплементиран коришћењем предложених функција за израчунавање мета података које се заснивају на доменском знању. Предложени мета подаци испуњавају критичне захтеве за коришћење у аутоматизованим системима за машинско учење.

Научни доприноси ове докторске дисертације су следећи:

- Извршена је евалуација постојећих скупова мета података који се користе у аутоматизованим системима за машинско учење и дат је предлог њихове

класификације са аспеката критичних захтева важних за проблем детекције аномалија и аутоматизованих система за машинско учење

- Предложен је нови скуп мета података заснованих на доменском знању који пружа боље карактеристике са аспекта критичних захтева важних за проблем одабира модела за детекцију аномалија
- Креиран и карактерисан је нови репозиторијум који обухвата скупове података релевантне за евалуацију предлога модела у аутоматизованим системима за машинско учење на основу доступних података у отвореној литератури и индустрији, који је допуњен са мета подацима заснованим на доменском знању
- Креиран је софтверски систем за одабир модела за детекцију аномалија на основу мета података, који је погодан за коришћење у аутоматизованим системима за машинско учење, који је проширив новим скуповима мета података, алгоритмима за детекцију аномалија и функцијама за рачунање удаљености између скупова података
- Дизајнирани су експерименти погодни за мерење перформанси предлога модела за детекцију аномалија на основу предложеног скупа мета података
- Извршена је евалуација перформанси креираног софтверског система са становишта различитих локација извршавања и архитектура

4.2. Критичка анализа резултата истраживања

Увидом у докторску дисертацију, полазне хипотезе, реализоване одлуке и добијене резултате, Комисија констатује да је кандидат успешно одговорио на постављене изазове и да резултати оправдавају дефинисане полазне хипотезе. Овај рад предлаже нови скуп мета података заснованих на доменском знању који може да се користи за одабир алгоритма за детекцију аномалија у аутоматизованим системима за машинско учење. Анализом постојећих мета података и увођењем критичних захтева везаних за скалабилност и перформансе, показано је да постоји потреба за дефинисањем новог скупа мета података у домену мета учења за одабир алгоритама. Резултати експеримената су показали да предложено решење постиже тачност од 87%, док постојећа решења постижу тачност од 73%, и при томе има значајно мању комплексност и доследно испуњава дефинисане критичне захтеве над креираним репозиторијумом који се састоји од 63 скупа података заступљених у индустрији. Предложено решење је поређено са постојећим решењима за различите типове података, различите домене података и различите локалитете аномалија у подацима. Решење постиже исте или боље резултате у поређењу са постојећим решењима за све типове аномалија, осим за локалне аномалије где само мета подаци засновани на теорији информација постижу за 9% боље резултате. Решење постиже исте или боље резултате у поређењу са постојећим решењима за све домене података осим за транспорт, где комплетан скуп мета података постиже за 25% боље резултате. Дефинисани скуп мета података је заснован искључиво на доменском знању и простим мета подацима, што испуњава критичне захтеве за коришћење у аутоматизованим системима за машинско учење. Показано је да доменски експерт може да процени предложене мета податке са нерастућом функцијом грешке искључиво у ситуацијама када постоји значајан број скупова података у систему и када грешка није већа од 50%. Такође, показано је да су функције за мерење удаљености засноване на трансформацијама података у мањи димензионални простор погодне за мерење сличности између скупова података коришћењем предложеног скупа мета података.

4.3. Верификација научних доприноса

У наставку је дат списак радова који су у непосредној вези са докторском дисертацијом:

Категорија M21:

Milos Kotlar, Marija Punt, Zaharije Radivojevic, Milos Cvetanovic, Veljko Milutinovic, „Novel

Meta-Features for Automated Machine Learning Model Selection in Anomaly Detection”, *IEEE Access*, vol. 9, pp. 89675-89687, ISSN: 2169-3536, DOI: 10.1109/ACCESS.2021.3090936, 2021. (IF 2019 = 3.745).

Категорија M22:

Milos Kotlar, Dragan Bojic, Marija Punt, Veljko Milutinovic, „Survey of deployment locations and underlying hardware architectures for contemporary deep neural networks”, *International Journal of Distributed Sensor Networks*, vol. 8, ISSN: 15501329, DOI: 10.1177/1550147719868669, 2019. (IF 2019 = 1.787).

5. ЗАКЉУЧАК И ПРЕДЛОГ

Комисија констатује да докторска дисертација кандидата Милоша Котлара испуњава све законске, суштинске и формалне услове предвођене Законом о високом образовању и Статутом и Правилником о докторским студијама Електротехничког факултета у Београду.

Докторска дисертација предлаже нови скуп мета података заснованих на доменском знању који може да се користи за одабир алгоритма за детекцију аномалија у аутоматизованим системима за машинско учење. Дефинисани скуп мета података је заснован искључиво на доменском знању и простим мета подацима, што испуњава критичне захтеве за коришћење у аутоматизованим системима за машинско учење. Кандидат Милош Котлар је у оквиру докторске дисертације извршио детаљну анализу постојећих решења, њихових предности и недостатака и дао предлог новог решења које постиже исте или боље резултате за аспекте од интереса. Главни допринос ове докторске дисертације представља нови скуп мета података заснован на доменском знању који је намењен за одабир алгоритма у аутоматизованим системима за детекцију аномалија.

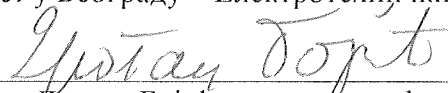
Имајући у виду наведено, Комисије предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да се докторска дисертација под називом **Детекција аномалија коришћењем мета података у аутоматизованим системима за машинско учење** кандидата Милоша Котлара прихвати, изложи на увид јавности и упуту на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

У Београду 28.03.2022.

ЧЛАНОВИ КОМИСИЈЕ



др Милош Цветановић, ванредни професор
Универзитет у Београду – Електротехнички факултет



др Драган Бојић, редовни професор
Универзитет у Београду – Електротехнички факултет



др Синиша Влајић, редовни професор
Универзитет у Београду – факултет Организационих наука



др Горан Квашчев, ванредни професор
Универзитет у Београду – Електротехнички факултет