

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Ивана Стојковића

Одлуком Наставно-научног већа бр. 5056/11-3 од 25. јануара 2018. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Ивана Стојковића под насловом

"Примена функционалних норми за регуларизацију рангирања над темпоралним подацима"

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

Иван Стојковић је уписао докторске студије на Електротехничком Факултету, Универзитета у Београду у пролећном семестру школске 2011/12 године.

Кандидат је тему докторске дисертације под радним називом „Структурирани приступ анализи комплексних биомедицинских система и сигнала“ пријавио Комисији за студије трећег степена на Електротехничком факултету, Универзитета у Београду, 01.06.2017. године и за ментора са Београдског Универзитета предложио проф. др Бранка Ковачевића. У складу са билатералним програмом докторских студија између Универзитета у Београду и Темпле Универзитета у Филаделфији, САД, ментор са стране Темпле Универзитета је проф. др Зоран Обрадовић.

Комисија за студије трећег степена разматрала је, на својој седници одржаној 06.06.2017. године, предлог теме за израду докторске дисертације и упутила предлог Комисије о оцени подобности теме и кандидата Наставно-научном већу на усвајање.

Наставно-научно веће Електротехничког факултета, на својој седници одржаној 20.06.2017. године (број одлуке 5056/11-1 од 20.06.2017), именovalo је Комисију за оцену услова и прихватање теме докторске дисертације у саставу: редовни професор др Жељко Ђуровић, редовни професор у пензији др Мирјана Поповић и доцент др Милош Вујисић.

На седници Наставно-научног већа Електротехничког факултета, одржаној 07.11.2017. године, усвојен је извештај Комисије за оцену услова и прихватање теме докторске дисертације (број одлуке 5056/11-2 од 07.11.2017) под новим (измењеним) насловом „Примена функционалних норми за регуларизацију рангирања над темпоралним подацима“.

Веће научних области техничких наука Универзитета у Београду је на седници од 27.11.2017. године дало сагласност на предлог теме докторске дисертације под насловом „Примена функционалних норми за регуларизацију рангирања над темпоралним подацима“ (број одлуке 61206-4779/2-17).

Кандидат је предао докторску дисертацију на преглед и оцену 28.12.2017. године. Комисија за студије трећег степена потврдила је на својој седници одржаној 10.01.2018. године испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације.

Наставно-научно веће Факултета је на својој седници, одржаној 25.01.2018. године именovalo Комисију за преглед и оцену докторске дисертације у саставу: редовни професор др Бранко Ковачевић (Електротехнички факултет), редовни професор др Жељко Ђуровић (Електротехнички факултет), редовни професор др Зоран Обрадовић (College of Science and Technology, Temple University), редовни професор др Слободан Вучетић (College of Science and Technology, Temple University), ванредни професор Kai Zhang (College of Science and Technology, Temple University), и редовни професор др Carmen Sapienza (Lewis Katz School of Medicine, Temple University), број одлуке 5056/11-3 од 25.01.2018. године.

На основу одлуке Наставно–научног већа бр. 545/2 од 13.03.2012. године, Студијски програм је започео у пролећном семестру школске 2011/2012, па се рок за завршетак докторских академских студија рачуна од почетка тог семестра, сагласно Статуту Универзитета у Београду и Статуту Електротехничког факултета. По истеку законског рока за завршетак докторских студија, на захтев студента, одобрено је продужење рока за завршетак студија за два семестра, сагласно Статуту Универзитета у Београду и Статуту Електротехничког факултета.

1.2. Научна област дисертације

Докторска дисертација припада техничким наукама и ужим научним областима машинског учења и анализе података, за које су Електротехнички факултет Универзитета у Београду и Колеџ за Науку и Технологију, Темпле Универзитета матични.

Ментор докторског рада са Електротехничког факултета Универзитета у Београду, проф. др Бранко Ковачевић је изабран у звање редовног професора за научне области обраде сигнала и управљања системима. Проф. др Зоран Обрадовић, ментор са Колеџа за Науку и Технологију, Темпле Универзитета у Филаделфији је изабран у звање редовног професора за научне области компјутерских наука и информационих система. Оба ментора су аутори великог броја научних радова у истакнутим међународним часописима.

1.3. Биографски подаци о кандидату

Иван Стојковић је рођен 11.04.1987. године у Београду, општина Савски венац. Основне студије је завршио на Електротехничком факултету, смер Сигнали и Системи, са просечном оценом 8.31. Дипломирао 2010. године, успешно одбранивши дипломски рад на тему: ”Симулатор робота SCARA конфигурације“ код професора Вељка Поткоњака. Мастер студије је завршио на истом смеру 2011., одбранивши мастер рад: “Управљање и навигација мобилних робота у непознатом окружењу” код професора Жељка Ђуровића. Докторске студије је уписао 2011/2012. године на модулу Управљање системима и обрада сигнала, и положио је све испите предвиђене Наставним планом и програмом модула са просечном оценом 10. Од јануара 2012. до августа 2014. године је био запослен као истраживач у Лабораторији за роботiku, при Институту Михајло Пупин у Београду.

Од фебруара 2014. године ангажован је као истраживач у Центру за Анализу Података и Биомедицинску Информатику на Темпле Универзитету у Филаделфији, где је учествовао на пројектима финансираним од стране Defense Advanced Research Projects Agency (DARPA):

1. “Graph-theoretic Research on Algorithms and the Phenomenology of Social networks (GRAPHS)” AFOSR award number FA 9550-12-1-0406
2. “Dialysis-Like Therapeutics (DLT)” SSC Pacific grant No. 66001-11-1-4183
3. “Technologies for Host Resilience (THoR)” under Contract No. W911NF-16-C-0050

У оквиру тих пројеката је сарађивао са реномираним стручњацима из престижних институција и стекао искуство рада у областима машинског учења, вештачке интелигенције и анализе података, претежно на апликацијама из биомедицинског домена.

У јесењем семестру 2014. године уписује докторске студије на Одсеку за Рачунарске Науке и Информатику, при Колеџу за Науку и Технологију, на Темпл Универзитету. Докторске студије Ивана Стојковића на Београдском Универзитету и на Темпле Универзитету су координисане уговором о дуалним докторским студијама између два универзитета.

Лета 2017. године је похађао праксу у Yahoo! Research Labs, на позицији Data Science практиканта, где је радио на развијању статистичких модела и анализу хетерогених података са више извора, у циљу идентификације прилика за апликацију рекомендер система.

Иван Стојковић је ко-аутор шест журналских радова, три поглавља у монографским збиркама и дванаест радова презентованих на међународним и домаћим конференцијама. Поменути радови се баве широким спектром проблема, од метода рангирања и пробабилистичких графичких модела, преко оптимизације терапије и биомедицинских апликација, па до управљања мобилним роботима и мултиагент системима.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Дисертација под насловом “Примена функционалних норми за регуларизацију рангирања над темпоралним подацима” написана је на 74 стране на енглеском језику, због билатералног докторског програма са Колеџом за Науку и Технологију, Темпле Универзитета у Филаделфији. Дисертација је организована у следећих седам поглавља: 1. Увод; 2. Релевантне студије; 3. Методологија: Мулти-проблем формулација; 4. Методологија: Високо-димензиона формулација; 5. Емпиријска евалуација: Мулти-проблем формулација; 6. Емпиријска евалуација: Високо-димензиона формулација; 7. Закључак. На почетку дисертације дати су Састав комисије, Резиме на енглеском, и на српском језику, Посвета, Захвалница, Листа акронима и Садржај. Иза последњег поглавља дата су два Додатака А) Подаци за Респираторне Вирусне Инфекције, и Б) Анализа Стабилности Селектованих Обележја, праћена списком коришћене литературе са 92 референце наведене по абecedном редоследу презимена првих аутора и Биографија кандидата.

2.2. Кратак приказ појединачних поглавља

Уводно поглавље састоји се од три дела. Први део описује мотивацију истраживања која се спроводе у тези, као и предмет и важност истраживања. Фокус истраживања је на аутоматском учењу функционалног мапирања из обсервабилних особина у циљну особину од интереса која је латентна (тј. није лако мерљива). У таквом случају када циљне лабеле нису директно доступне за обуку параметарских модела, може се користити неки други облик супервизованих информација, као што је информација о делимичном редоследу међу примерима (један пацијент изгледа здравије од другог). Важност описане теме произилази из

његове распрострањености, пошто многе карактеристике од интереса (које би неко желео да измери) су суштински латентне, као што су здравље, интелигенција или ризик неког улагања. Споменуте су савремене методе за решавање таквог проблема и наведени су њихови недостаци.

У другом делу уводног поглавља истакнути су додатни изазови који произилазе из процеса прикупљања података, а начини за њихово решавање су предложени и укратко разматрани. Трећи део уводног поглавља приказује структуру тезе и кратак опис тема о којима ће се дискутовати у наредним поглављима.

Друго поглавље представља савремене технике и описује релевантне студије у контексту истраживања спроведених у докторској дисертацији. Прегледане су савремене методе за учење функција оцењивања засноване на класификацији, регресији и рангирању. Дискутовано је и коришћење функционалних норми за регуларизацију више-димензионалних и мулти-проблема. Поред тога, проксимални алгоритми су покривени као популарни алат за оптимизацију претходно дискутованих приступа.

Треће поглавље је методолошко поглавље које описује предложену Мулти-проблем формулацију регуларизованог рангирања за темпоралне податке. Комплетно извођење формулације модела заједно са прилагођеним алгоритмом оптимизације су детаљно описане у овом поглављу.

Четврто поглавље се бави високо-димензионалном формулацијом регуларизованог рангирања за темпоралне податке. Поглавље представља извођење високо-димензионалног модела са одговарајућим алгоритмом оптимизације за прилагођавање његових параметара на основу измерених података.

Пето поглавље даје резултате емпиријске процене Мулти-проблем модела. Приступ је евалуиран на синтетичким подацима и на две реалне апликације. Једна апликација је учење оцена из испита ученика основних школа, док је циљ друге апликације учење толеранције на респираторне вирусне инфекције код људи. Резултати сугеришу повећану тачност предвиђања предложеног приступа над алтернативама које су обучаване на појединачним задацима.

Шесто поглавље описује резултате емпиријске процене модела високо-димензионалног рангирања. Модел се прво процењује на сету интуитивних синтетичких примера, где су очигледне предности “проређених” модела у односу на “густе”. Одељак резултата наставља са проценом на реалним апликацијама, на подацима о експресији гена код Х3Н2 вирусне инфекције код људи, и на подацима о експресији гена пуноглаваца (жаба) приликом бактеријске инфекције. Примећена је унапређена тачност, као и робусност предложеног метода, у односу на алтернативне методе.

У последњем, седмом поглављу, коментарисан је научни допринос, резултати су сумирани и предлажу се даљи правци истраживања. Поред тога, у наставку су дата још два поглавља Додатака. Први описује темпоралне податке о људској вирусној инфекцији, који су коришћени за процену Мулти-проблем методе, а други Додатак даје додатне детаље о анализи стабилности одабира обележја код високо-димензионалне методе.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Пораст количине и разноврсности података који се сакупљају и обрађују ради стицања корисног увида, је задњих година довео до видљиве трансформације индустрије. Све је више производа и апликација које су засноване на коришћењу алгоритама машинског учења за извлачење образаца из прикупљених података. И у академској сфери је видљив сличан тренд повећаног интереса за развој алгоритама машинског учења, расте и број нових специјализованих конференција као и број радова на већ утврђеним конференцијама у области. Тако да је и тема кандидата у складу са тренутним трендом.

Са порастом броја и разноврсности апликација, појављују се и нови изазови. И иако су проблеми као што је учење из парцијалног поредка примера, или учење из високодимензионалних примера већ доста добро проучени појединачно, комбинације више таквих околности још увек представљају изазов. Оригиналност теме кандидата се налази у адресирању претходно непроучаваних комбинација као што је проблем рангирања над темпоралним подацима у високодимензионалним апликацијама, или проблем учења рангирања над темпоралним подацима из више повезаних проблема. Кандидат је те проблеме формализовао као линеарне моделе са регуларизацијом на бази функционалних норми и критеријумској функцији заснованој на максимизацији маргине. Поред предлагања нових формулација модела, кандидат је пружио и специјализоване алгоритме за ефикасно обучавање предложених модела.

3.2. Осврт на референтну и коришћену литературу

У овој докторској дисертацији, анализирана литература је обимна и релевантна. Наведено је укупно 92 библиографске референце, из области конвексне оптимизације, супервизованог учења и биомедицинског домена. Литература се претежно састоји од радова новијег датума, обзиром да су обрађивани проблеми последица скоријег технолошког напретка који је омогућио прикупљање података велике количине и разноврсности. То додатно потврђује актуелност теме и истраживања обухваћеног овом докторском тезом. Списак литературе укључује и релевантне радове које је кандидат публиковао као аутор или коаутор.

3.3. Опис и адекватност примењених научних метода

Истраживање у оквиру ове докторске дисертације је засновано на следећим експерименталним, теоријским и практичним методама:

- Методе за предпроцесирање података: (1) Интерполација података са циљем надокнађивања недостајућих података; (2) Нормализација (или стандардизација) података тако да се магнитуде променљивих доведу на исти (упоредиви) ред величине.
- Методе супервизованог учења: Параметри (линеарних) модела су обучавани тако да се минимизује одређена критеријумска функција која узима у обзир мерене податке.
- Методе селекције релевантних обележја: Код високо димензионалних проблема јавља се проблем слабо релевантних обележја који умањују генерализационе могућности модела. Због тога се примењују технике селекције обележја, како би се задржала само она обележја од високог значаја.
- Методе регуларизације: Да би се спречио (или ублажио) проблем преобучавања, уводе се додатне претпоставке о вредностима параметара кроз пенализовање одступања од номиналних карактеристика (на пример вредности треба да буду блиске нули).
- Методи конвексне оптимизације: За обучавање параметара (конвексних) модела користе се ефикасни алгоритми са јаким гаранцијама квалитета добијених решења.
- Методи за мерење перформанси: Генерализационе могућности обучених модела се испитују на претходно невиђеним тест подацима и мере стандардним мерама тачности предикција у односу на праве вредности.
- Статистичке методе: Статистичке методе се примењују за испитивање значајности или конзистентности добијених резултата.

3.4. Применљивост остварених резултата

Предложени објективи за моделирање су погодни за проблеме у којима је лакше добити информације о делимичном редоследу, него циљне лабеле за регресију или класификацију. Погодни примери су релевантност добијених резултата претраге на интернету, компетентности неког кандидата за посао или здравственог стања неког пацијента. Темпорални подаци и проблеми су такође свеprisутни и често је врло важно да апликација укључи временски аспект, као што је на пример улагање у тржиште акција. Према томе, велики број ситуација може се на врло прикладан начин моделирати као рангирање над темпоралним подацима. Дobar део таквих апликација такође ће имати велики број измерених обележја, за које би предложени високодимензионални модел био добар избор.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат Иван Стојковић је израдом ове докторске дисертације и публикацијом преко двадест научних радова у потпуности демонстрирао способности за самосталан научно-истраживачки рад. Кандидат је показао способност праћења и разумевања релевантне и актуелне литературе. Успешно је идентификовао недостатке досадашњих студија и предложио решења којима се превазилазе ограничења постојећих приступа. Тема дисертације је веома актуелна и мултидисциплинарна, при чему резултати дисертације имају потенцијал за практичне примене. Методе истраживања и приступи у реализацији су релевантни и иновативни. Остварени доприноси су оригинални и потврђују способности кандидата за самостални научно-истраживачки рад.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Научни доприноси ове докторске дисертације су следећи:

- Развијен је математички модел за учење функционалног мапирања из вискодимензионалних темпоралних обсервација и парцијално поређаних парова по циљној варијабли.
- Развијен је математички модел за учење функционалног мапирања из темпоралних обсервација и парцијално поређаних парова по циљној варијабли, за више различитих (али зависних) проблема.
- Развијен је оптимизациони алгоритам за обучавање параметара мулти-проблем модела за темпорално рангирање, заснован на проксималним алгоритмима
- Модел је примењен на проблем проучавања толеранције на инфекције код људи.

4.2. Критичка анализа резултата истраживања

Предности предложене методологије су могућност учења и у приликама кад променљива од интереса није доступна за директно мерење већ је могуће само упоредити примере међусобом. Тако се отвара могућност за учење функционалног мапирања које после може да се користи као мерни инструмент за квантификацију интензитета променљиве од интереса, што је практичан и чисто емпиријски начин употребе предложене методологије.

Употреба функционалних норми за регуларизацију које индукују проређени вектор параметара модела (код вишедимензионе формулације), може да служи као метода за селекцију обележја која су значајно корелисана са променљивом од интереса. Селектована обележја могу да сугеришу дубљу зависност са циљаном варијаблом (потенцијално узрочну), што захтева даља истраживања фокусирана на селектована обележја. Тако да се предложена

методологија може користити и са теоријске стране за генерисање потенцијалних хипотеза о везама између променљивих.

Још једна практична примена предложене (мулти проблем) методологије је употреба у ситуацијама када појединачни задаци имају премало примера за добру генерализацију, где би се кроз симултано обучавање више модела индиректно ојачали заједнички (дељени) образци у подацима.

4.3. Верификација научних доприноса

Научни доприноси докторске дисертације су верификовани следећим радовима:

Категорија M21:

1. **Stojkovic I.**, Obradovic Z.: Sparse Learning of the Disease Severity Score for High-Dimensional Data,- Complexity, vol. 2017, Article ID 7120691, pp. 1-11, 2017. DOI:10.1155/2017/7120691 ISSN: 1099-0526 (online) (2016 impact factor 4.621) (9/63 Multidisciplinary Sciences; 2/100 Mathematics, Interdisciplinary Applications)
2. **Stojkovic I.**, Ghalwash MF., Cao XH., Obradovic Z.: Effectiveness of Multiple Blood-Cleansing Interventions in Sepsis, Characterized in Rats, - Scientific Reports, Vol. 6 (24719), pp. 1-11, Nature Publishing Group, 2016. DOI: 10.1038/srep24719 ISSN: 2045-2322 (online) (2016 impact factor 4.259) (10/63 in Multidisciplinary Sciences)
3. Cao XH., **Stojkovic I.**, Obradovic Z.: A robust data scaling algorithm to improve classification accuracies in biomedical data, - BMC Bioinformatics, Vol. 17 (359), pp. 1-10, BioMed Central, 2016. DOI: 10.1186/s12859-016-1236-x ISSN:1471-2105 (online) (2016 impact factor 2.448) (10/57 in Mathematics & Computational Biology; 38/78 Biochemical Research Methods; 68/160 Biotechnology & Applied Microbiology)
4. Ghalwash MF., Cao XH., **Stojkovic I.**, Obradovic Z.: Structured feature selection using coordinate descent optimization, - BMC Bioinformatics, Vol. 17 (158), pp. 1-14, BioMed Central, 2016. DOI: 10.1186/s12859-016-0954-4 ISSN:1471-2105 (online) (2016 impact factor 2.448) (10/57 in Mathematics & Computational Biology; 38/78 Biochemical Research Methods; 68/160 Biotechnology & Applied Microbiology)

Категорија M33:

1. **Stojkovic I.**, Ghalwash MF., Obradovic Z.:Ranking Based Multitask Learning of Scoring Functions, - Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Skopje, Macedonia, September 2017
2. Pavlovski M., Zhou F., **Stojkovic I.**, Kocarev Lj., Obradovic Z.:Adaptive Skip-Train Structured Regression for Temporal Networks, - Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Skopje, Macedonia, September 2017
3. **Stojkovic I.**, Jelisavcic V., Milutinovic V., Obradovic Z.:Fast sparse Gaussian Markov Random Fields learning based on Cholesky factorization,- Proc. 26th International Joint Conference on Artificial Intelligence (IJCAI), Melbourne, Australia, August 2017
4. **Stojkovic I.**, Obradovic Z.:Predicting Sepsis Biomarker Progression under Therapy, -Proc. 30th IEEE International Symposium on Computer-Based Medical Systems, Thessaloniki, Greece, June, 2017
5. **Stojkovic I.**, Jelisavcic V., Milutinovic V., Obradovic Z.: Distance Based Modeling of Interactions in Structured Regression, - Proc. 25th International Joint Conference on Artificial Intelligence IJCAI-16, New York City, NY, USA, July 2016
6. Cao XH., **Stojkovic I.**, Obradovic Z.: Predicting sepsis severity from limited temporal observations, -Proc. 17th Discovery Science, Bled, Slovenia, October 2014

5. ЗАКЉУЧАК И ПРЕДЛОГ

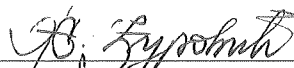
На основу свега изложеног, Комисија сматра да докторска дисертација кандидата Ивана Стојковића испуњава све законске, формалне и суштинске услове, као и све критеријуме који се уобичајено примењују приликом вредновања докторске дисертације. Узимајући у обзир наведене научне доприносе, у контексту предложених метода за регуларизовано рангирање над темпоралним подацима, њихову показану применљивост у биомедицинским апликацијама, као и зрелост и способност кандидата за самосталан научно истраживачки рад, Комисија сматра да докторска дисертација Ивана Стојковића садржи оригиналне научне доприносе у областима машинског учења и анализе података у биомедицинским апликацијама. Стога Комисија са задовољством предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да се докторска дисертација под називом "Примена функционалних норми за регуларизацију рангирања над темпоралним подацима" прихвати, изложи на увид јавности и упуту на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

Београд, 01.02.2018. године

ЧЛАНОВИ КОМИСИЈЕ



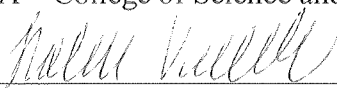
др Бранко Ковачевић, редовни професор
Универзитет у Београду – Електротехнички факултет



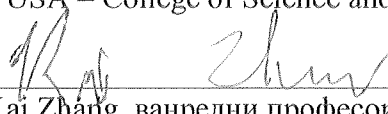
др Жељко Ђуровић, редовни професор
Универзитет у Београду – Електротехнички факултет



др Зоран Обрадовић, L.H. Carnell Professor
Temple University, USA – College of Science and Technology



др Слободан Вучећић, редовни професор
Temple University, USA – College of Science and Technology



др Kai Zhang, ванредни професор
Temple University, USA – College of Science and Technology



др Carmen Sapienza, редовни професор
Temple University, USA – Fox Chase Cancer Center

TO TEACHING-SCIENTIFIC COUNCIL

Subject: Report about completed doctoral dissertation of candidate Ivan Stojković

By the decision of Teaching-scientific council of the School of Electrical Engineering number 5056/11-3 from January, 25th, 2018, we are selected as Committee members for examination, evaluation and defense of doctoral dissertation of candidate Ivan Stojković with the title

"Functional Norm Regularization for Margin-Based Ranking on Temporal Data"

After examination of the submitted thesis and accompanying material and interview with the candidate, the Committee made the following

R E P O R T

1. INTRODUCTION

1.1. Chronology of approval and dissertation work

Ivan Stojković enrolled into doctoral program at School of Electrical Engineering at University of Belgrade in Spring semester of school 2011/2012 year.

The candidate has submitted the thesis topic with the working title „Structured approach to analysis of complex biomedical systems and signals“ to the Commission for the third-level studies at the School of Electrical Engineering, University of Belgrade on June, 1st, 2017, and proposed prof. dr Branko Kovačević as a supervisor from the Belgrade University side. In accordance with the bilateral program of doctoral studies between the University of Belgrade and Temple University in Philadelphia, USA, supervisor from the Temple University side is prof. dr Zoran Obradović.

Commission for the third-level studies considered the proposal of the thesis topic for dissertation work on its meeting organized on June, 6th 2017 and forwarded the proposal of the Committee for evaluation of the suitability of the thesis topic and candidate to the Teaching-scientific council.

Teaching-scientific council of the School of Electrical Engineering on its meeting organized on June, 20th 2017 (decision number 5056/11-1 from June, 20th 2017) declared the Committee for evaluation of the conditions and acceptance of the doctoral dissertation topic, composed from: prof. dr Željko Đurović, (retired) prof. dr Mirjana Popović and assistant prof. dr Miloš Vujisić.

On the meeting of Teaching-scientific council of the School of Electrical Engineering, organized on November, 7th 2017, the report of the Committee for evaluation of the conditions and acceptance of the doctoral dissertation topic has been accepted (decision number 5056/11-2 from November, 7th 2017) under the new (changed) title „Functional Norm Regularization for Margin-Based Ranking on Temporal Data“.

Council of scientific areas of technical sciences of the University of Belgrade, on its meeting organized on November, 27th 2017 gave the consent to the proposal of doctoral dissertation topic with the title „Functional Norm Regularization for Margin-Based Ranking on Temporal Data“ (decision number 61206-4779/2-17).

The candidate has submitted the doctoral dissertation for examination and evaluation on December, 28th 2017.

Commission for the third-level studies, on its meeting organized on January, 10th 2018, has confirmed the fulfillment of all the necessary conditions for proposal submission to the Teaching-scientific council of the School of Electrical Engineering for forming the Committee for examination and evaluation of doctoral dissertation.

Teaching-scientific council of the School of Electrical Engineering, on its meetings organized on January, 25th, 2018, declared the Committee for examination and evaluation of doctoral dissertation, composed from: prof. dr Branko Kovačević (School of Electrical Engineering, University of Belgrade), prof. dr Željko Đurović (School of Electrical Engineering, University of Belgrade), prof. dr Zoran Obradović (College of Science and Technology, Temple University), prof. dr Slobodan Vučetić (College of Science and Technology, Temple University), associate professor dr Kai Zhang (College of Science and Technology, Temple University), and prof. dr Carmen Sapienza (Lewis Katz School of Medicine, Temple University), decision number 5056/11-3 from January 25th, 2018.

Based on the decision of the Teaching-Scientific Council no. 545/2 from 13/03/2012, the study program started in the spring semester of the academic year 2011/2012, so the deadline for completion of doctoral academic studies is calculated from the beginning of that semester, in accordance with the Statute of the University of Belgrade and the Statute of the Faculty of Electrical Engineering. Upon expiration of the deadline for completion of doctoral studies, at the request of the student, an extension of the deadline for completion of studies for two semesters was granted, in accordance with the Statute of the University of Belgrade and the Statute of the Faculty of Electrical Engineering.

1.2. Scientific areas of doctoral dissertation

The doctoral dissertation belongs to Technical sciences and research areas of machine learning and data analytics, for which the School of Electrical Engineering, University of Belgrade and the College of Science and Technology at Temple University are specialized.

Thesis supervisor from the School of Electrical Engineering, University of Belgrade, prof. dr Branko Kovačević is full professor for scientific areas of signal processing and control systems. Thesis supervisor from the College of Science and Technology, Temple University, prof. dr Zoran Obradović is full professor for scientific areas of computer science and information systems. Both supervisors are authors of large number of scientific papers in the international journals.

1.3. Biographical data

Ivan Stojković was born on April 11th, 1987. in Belgrade, municipality Savski Venac. He completed his undergraduate studies at the School of Electrical Engineering, Department of Signals and Systems, with an average grade of 8.31. He graduated in 2010, successfully defending his thesis on the topic: "Simulator of the SCARA Configuration Robot", guided by Professor Veljko Potkonjak. At the same department, he completed the Master studies in 2011, defending master thesis: "Control and Navigation of Mobile Robots in an Unknown Environment" under supervision of Professor Željko Đurović. He enrolled into doctoral studies in year 2011/2012 on the module for Control systems and signal processing, and passed all the exams stipulated by the Curriculum and module program with average grade of 10. From January 2012 to August 2014, he was employed as a researcher at the Laboratory for Robotics at the Mihajlo Pupin Institute in Belgrade.

Since February 2014, he has been engaged as a researcher at the Center for Data Analytics and Biomedical Informatics at Temple University in Philadelphia, where he participated in projects funded by the Defense Advanced Research Projects Agency (DARPA):

1. "Graph-theoretical Research on Algorithms and the Phenomenology of Social Networking (GRAPHS)" AFOSR award number FA 9550-12-1-0406
2. "Dialysis-Like Therapeutics (DLT)" SSC Pacific grant No. 66001-11-1-4183
3. "Technologies for Host Resilience (THoR)" under Contract No. W911NF-16-C-0050

Within these projects he collaborated with renowned experts from prestigious institutions and gained experience in the fields of machine learning, artificial intelligence and data analysis, predominantly on applications from a biomedical domain.

In the fall of 2014, he enrolled in doctoral studies at the Department for Computer and Information Sciences, at the College of Science and Technology, at Temple University. Doctoral studies of Ivan Stojković at Belgrade University and at Temple University are coordinated by a dual doctoral studies agreement between the two universities.

In summer of 2017, he interned at Yahoo! Research Labs, on a position of Data Science intern, where he worked on developing statistical models and analyzing heterogeneous data from multiple sources, in order to identify the opportunities for the application of the recommender systems.

Ivan Stojković is author or co-author of six journal papers, three chapters in monographic collections and twelve papers presented at international and domestic conferences. Mentioned articles deal with a wide range of problems ranging from ranking methods and probabilistic graphic models, over therapy optimization and biomedical applications, to the control of mobile robots and multiagent systems.

2. DESCRIPTION OF DISSERTATION

2.1. Dissertation content

Dissertation with the title "Functional Norm Regularization for Margin-Based Ranking on Temporal Data" is written on 74 pages in English language, due to the bilateral doctoral program with the College of Science and Technology at Temple University in Philadelphia. Dissertation is organized in the following seven chapters: 1. Introduction; 2. Related Work; 3. Methodology: Multi-task Formulation; 4. Methodology: High-dimensional Formulation; 5. Results: Multi-task Framework; 6. Results: High-dimensional Framework 7. Conclusions. At the beginning of dissertation, the Committee, the Abstracts in English and Serbian languages, Dedication, Acknowledgement, List of acronyms and Contents are written. After the last chapter, there are two Appendices: A) Respiratory Viral Infection Data, and B) Feature Selection Stability, followed by the Bibliography of 92 references listed by the surname of the first author and Biography of the candidate.

2.2. A brief of overview of thesis chapters

The introductory chapter consists of three parts. The first part describes the motivation for research carried out in the thesis, as well as the subject and importance of the research. The research focus is on automatic learning of functional mapping from observable features into the property of interest which is latent (i.e. not easily observable). In such case when the target labels are not directly available for training of parametric models, some other form of supervised information might be used instead, like information on partial order among the examples (this patient appears healthier than the other one). Importance of the described topic arises from its prevalence, as many of the

interesting characteristics that one would like to quantify are essentially latent, like health, intelligence, or risk of some investment. Contemporary methods to deal with such problem are mentioned and their shortcomings are outlined.

In the second part of the introductory chapter, additional challenges arising from the data gathering process are stressed out and approaches to address them are proposed and briefly discussed.

The third part of the introductory chapter provides the thesis structure and a brief description of topics that will be discussed in the following chapters.

The second chapter presents the state of the art and describes the relevant studies in the context of research conducted in the doctoral dissertation. A contemporary methods to scoring function learning based on the classification, regression and ranking approaches are reviewed. Use of the functional norm regularization to tackle high-dimensional and multitask problems are discussed too. In addition, proximal algorithms are covered as a popular optimization tool for previously discussed approaches.

The third chapter is a methodological chapter which describes the proposed Multi-task formulation of the norm regularized ranking for the temporal data. Complete derivation of the model formulation along with the custom optimization algorithm are detailed here.

The fourth chapter deals with the High-dimensional formulation of the norm regularized ranking for the temporal data. The chapter present the derivation of the High-dimensional model with the appropriate optimization algorithm for fitting its parameters based on the data.

The fifth chapter gives results from the empirical evaluation of the Multi-task ranking model. The approach is evaluated on a synthetic data and in two real-world applications. One real application is learning exam scores of elementary school pupils, while the objective of the second application is learning the tolerance to respiratory viral infections in humans. The results suggests increased prediction accuracy of the proposed approach over the alternatives that are trained on individual tasks.

The sixth chapter describes the results from empirical evaluation of High-dimensional ranking model. The model is first evaluated on a set of intuitive synthetic examples, where the advantages of sparse framework over the non-sparse one are apparent. The results section continues with the assessment on a real-life applications, a gene expression dataset of H3N2 viral infection responses and gene expression data of (frog) tadpole bacterial infection. Efficacy, as well as the robustness of the proposed method, are compared favorably against multiple alternative methods.

In the last, seventh chapter, achieved scientific contributions and results are summarized and the further research directions are proposed. In addition, two more Appendices chapters are provided afterwards. First one is detailing the temporal human viral infection data used for evaluation of the Multi-task method, and the second appendix giving additional details on feature selection stability analysis of High-dimensional method.

3. DISSERTATION EVALUATION

3.1. Contemporaneity and originality

An increase in the quantity and diversity of the collected and processed data, with aim of getting useful insights, has led to the visible transformation of the industry in recent years. There are increasingly many products and applications that are based on the use of machine learning algorithms to extract patterns from collected data. Even in the academic sphere there is a similar trend of increased interest in the development of machine learning algorithms, and the number of new specialized conferences as well as the number of papers submitted at already established conferences in the field, is growing. So the subject of the candidate's thesis is well aligned with the current trends.

With the increase in the number and diversity of applications, new challenges arise. And although problems such as learning from the partially ordered examples, or learning from high-dimensional cases have already been sufficiently well studied individually, combinations of several such circumstances still pose a challenge. The originality of the candidate's work is in addressing previously uninvestigated combinations, such as the problem of ranking over temporal data in high-dimensional applications, or the problem of learning ranking over temporal data from multiple related problems. The candidate formalized these problems as linear model with regularization based on functional norms and a criterion function based on the maximization of the margin. In addition to proposing new model formulations, the candidate also provided specialized algorithms for effective training of the proposed models.

3.2. Review of the used literature

In this doctoral dissertation, the analyzed literature is extensive and relevant. A total of 92 bibliographic references have been cited, from areas of convex optimization, supervised learning and biomedical domain. The literature is mostly composed of the recent studies, given that concerned problems are a result of the recent technological advances that enabled the collection of data of large quantities and diversity. This fact additionally confirms the actuality of the topic and research in this doctoral thesis. The list of literature includes papers related to the thesis research that the candidate has published as an author or co-author.

3.3. Description and adequacy of applied scientific methods

Research in the scope of this doctoral dissertation is based on the following experimental, theoretical and practical methods:

- Methods for data preprocessing: (1) Interpolation of data for the purpose of compensating for missing data; (2) Normalization (or standardization) of data so that the magnitudes of the variables are brought to the same (comparable) order of magnitude.
- Supervised Learning Methods: The parameters of (linear) models are trained to minimize a particular criterion function that takes into account the measured data.
- Methods of selection of relevant features: In the case of highly dimensional problems, the problem of poorly relevant features arises which diminish the generalization possibilities of the model. Therefore, selection techniques are applied in order to retain only those features of high importance.
- Regularization methods: In order to prevent (or mitigate) the problem of overfitting, additional assumptions about parameter values are introduced by penalizing deviations from nominal characteristics (for example, values should be close to zero).
- Convex optimization methods: Effective algorithms with strong guarantees on the quality of the solutions obtained are used to train parameters of (convex) models.
- Performance Measurement Methods: The generalization abilities of the fitted models are examined on previously unseen test data and are measured by standard measurements of the accuracy of the predictions with respect to the true values.
- Statistical methods: Statistical methods are used to test the significance or consistency of the obtained results.

3.4. Applicability of the dissertation results

The proposed modeling objectives are suited for problems where information about partial order is easier to obtain than the regression or classification labels. Examples being relevance of the retrieved results from web search, competence of some candidate for a job, or healthiness of some patient. Temporal data and problems are also ubiquitous, and often it is vital for the application to

include the temporal aspect, like in making investments in stock market. Therefore, a number of situations can be appropriately modeled as ranking on the temporal data. Many of such applications will also have a huge number of measured features, for which the proposed high-dimensional framework would be a good fit.

3.5. Evaluation of the candidate's achieved capabilities for independent scientific work

Candidate Ivan Stojković has fully demonstrated the ability for independent scientific work by completion of this doctoral dissertation and publication of more than twenty scientific papers. The candidate demonstrated the ability to follow and understand the relevant and current literature. He successfully identified the shortcomings of the previous studies, and proposed solutions that overcome the limitations of existing approaches. The topic of the dissertation is very actual and multidisciplinary, while the dissertation results have a potential for practical applications. Research methods and developed approaches are relevant and innovative. The achieved contributions are original and confirm the candidate's ability for independent scientific research work.

4. ACHIEVED SCIENTIFIC CONTRIBUTION

4.1. Overview of the achieved scientific contributions

Scientific contributions of this doctoral thesis are the following:

- Mathematical model was developed for learning functional mapping from high-dimensional temporal observations and partially ordered pairs by the target variable.
- Mathematical model was developed for learning functional mapping from temporal observations and partially ordered pairs by the target variable, for multiple distinct (but dependent) problems.
- Optimization algorithm was developed for training the parameters of the multi-problem model for temporal ranking, based on proximal algorithms
- The model was applied to the problem of studying tolerance to infection in humans.

4.2. Critical analysis of the research results

The advantages of the proposed methodology are the ability to learn even in cases when the variable of interest is not available for direct measurement, but it is only possible to compare examples with each other. This opens up the possibility of learning functional mapping, which can then be used as a measuring instrument for quantifying intensity of variables of interest, which is a practical and purely empirical way of using the proposed methodology.

The use of functional norms for regularization which induce a sparse vector of model parameters (in a high-dimensional formulation) can serve as a method for selecting features that are significantly correlated with a variable of interest. Selected features may suggest a deeper dependency on the target variable (potentially causal), which requires further research focused on the selected features. So the proposed methodology can also be used on the theoretical side, to generate prospective hypotheses about the relationships between variables.

Another practical application of the proposed (multi-problem) methodology is the use in situations where individual tasks have too few examples for a good generalization, whereby simultaneous training of several models would indirectly reinforce common (shared) data patterns.

4.3. Verification of scientific contributions

Scientific contributions of this doctoral dissertation are verified based on the following publications:

Category M21:

1. **Stojkovic I.**, Obradovic Z.: Sparse Learning of the Disease Severity Score for High-Dimensional Data,- Complexity, vol. 2017, Article ID 7120691, pp. 1-11, 2017. DOI:10.1155/2017/7120691 ISSN: 1099-0526 (online) (2016 impact factor 4.621) (9/63 Multidisciplinary Sciences; 2/100 Mathematics, Interdisciplinary Applications)
2. **Stojkovic I.**, Ghalwash MF., Cao XH., Obradovic Z.: Effectiveness of Multiple Blood-Cleansing Interventions in Sepsis, Characterized in Rats, - Scientific Reports, Vol. 6 (24719), pp. 1-11, Nature Publishing Group, 2016. DOI: 10.1038/srep24719 ISSN: 2045-2322 (online) (2016 impact factor 4.259) (10/63 in Multidisciplinary Sciences)
3. Cao XH., **Stojkovic I.**, Obradovic Z.: A robust data scaling algorithm to improve classification accuracies in biomedical data, - BMC Bioinformatics, Vol. 17 (359), pp. 1-10, BioMed Central, 2016. DOI: 10.1186/s12859-016-1236-x ISSN:1471-2105 (online) (2016 impact factor 2.448) (10/57 in Mathematics & Computational Biology; 38/78 Biochemical Research Methods; 68/160 Biotechnology & Applied Microbiology)
4. Ghalwash MF., Cao XH., **Stojkovic I.**, Obradovic Z.: Structured feature selection using coordinate descent optimization, - BMC Bioinformatics, Vol. 17 (158), pp. 1-14, BioMed Central, 2016. DOI: 10.1186/s12859-016-0954-4 ISSN:1471-2105 (online) (2016 impact factor 2.448) (10/57 in Mathematics & Computational Biology; 38/78 Biochemical Research Methods; 68/160 Biotechnology & Applied Microbiology)

Category M33:

1. **Stojkovic I.**, Ghalwash MF., Obradovic Z.:Ranking Based Multitask Learning of Scoring Functions, - Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Skopje, Macedonia, September 2017
2. Pavlovski M., Zhou F., **Stojkovic I.**, Kocarev Lj., Obradovic Z.:Adaptive Skip-Train Structured Regression for Temporal Networks, - Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Skopje, Macedonia, September 2017
3. **Stojkovic I.**, Jelisavcic V., Milutinovic V., Obradovic Z.:Fast sparse Gaussian Markov Random Fields learning based on Cholesky factorization,- Proc. 26th International Joint Conference on Artificial Intelligence (IJCAI), Melbourne, Australia, August 2017
4. **Stojkovic I.**, Obradovic Z.:Predicting Sepsis Biomarker Progression under Therapy, -Proc. 30th IEEE International Symposium on Computer-Based Medical Systems, Thessaloniki, Greece, June, 2017
5. **Stojkovic I.**, Jelisavcic V., Milutinovic V., Obradovic Z.: Distance Based Modeling of Interactions in Structured Regression, - Proc. 25th International Joint Conference on Artificial Intelligence IJCAI-16, New York City, NY, USA, July 2016
6. Cao XH., **Stojkovic I.**, Obradovic Z.: Predicting sepsis severity from limited temporal observations, -Proc. 17th Discovery Science, Bled, Slovenia, October 2014

5. CONCLUSION AND PROPOSAL

Based on all the items mentioned above, the Committee considers that doctoral dissertation of the candidate Ivan Stojković fully meets all legal, formal and essential requirements, as well as all the criteria that are usually applied when evaluating the doctoral dissertation. Taking into account specified (expected and attained) scientific contributions, in the context of the proposed functional norm regularized ranking approaches for temporal data, its applicability in biomedical domain, and demonstrated ability of the doctoral candidate to conduct independent research, the Committee considers that the dissertation of Ivan Stojković contains original scientific contributions in the fields of machine learning and data analysis in biomedical applications. Therefore, the Committee is pleased to propose to the Teaching-Scientific Council of the Faculty of Electrical Engineering, University of Belgrade to accept the doctoral dissertation entitled "Functional Norm Regularization for Margin-Based Ranking on Temporal Data", put it on the public display and forward it for the final acceptance to the Council of scientific areas of technical sciences of University of Belgrade.

Belgrade, February, 1st 2018.

COMMITTEE MEMBERS



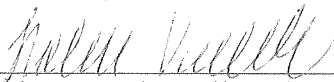
dr Branko Kovačević, Full professor
University of Belgrade – School of Electrical Engineering



dr Željko Đurović, Full professor
University of Belgrade – School of Electrical Engineering



dr Zoran Obradović, L.H. Carnell Professor
Temple University USA – College of Science and Technology



dr Slobodan Vučetić, Full professor
Temple University USA – College of Science and Technology



dr Kai Zhang, Associate professor
Temple University USA – College of Science and Technology



dr Carmen Sapienza, Full professor
Temple University USA – Fox Chase Cancer Center