



УНИВЕРЗИТЕТ У БЕОГРАДУ
Факултет организационих наука

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Подобност теме и кандидаткиње Тијане Вујичић за израду докторске дисертације

Одлуком Наставно-научног већа ФОН-а бр. 3/81-3 од 28.06.2017. године, именовани смо за чланове Комисије за оцену подобности теме и кандидаткиње Тијане Вујичић за израду докторске дисертације и научне заснованости теме „Софтверски алат за испитивање алгоритама структурне регресије базиране на *GCRF* моделу“.

На основу материјала приложеног уз захтев кандидаткиње, Комисија подноси следећи

РЕФЕРАТ

1. Подаци о кандидату

1.1 Биографски подаци

Тијана Вујичић рођена је 20. октобра 1990. године у Никшићу, где је завршила основну школу „Ратко Жарић“ и гимназију „Стојан Церовић“ са одличним успехом. Основне и специјалистичке студије завршила је на Факултету за информационе технологије Универзитета „Медитеран“ у Подгорици. Дипломирала је 2013. године на смеру Софтверско инжењерство, одбраном дипломског рада на тему „Моделовање података засновано на временски базираним подацима помоћу *Datomic NoSQL*-а“. Магистрирала је 2015. године на Факултету организационих наука Универзитета у Београду, на студијском програму Софтверско инжењерство и рачунарске науке, модул Софтверско инжењерство, одбраном тезе на тему „Коришћење неуронских мрежа за предвиђање употребљивости података“. Ментор за завршни рад на основним и мастер студијама био је др. Драган Ђурић. Основне, специјалистичке и мастер студије завршила је са просечном оценом 10,00 (десет, 0/100).

Од септембра 2013. године ради као асистент на Факултету за информационе технологије Универзитета „Медитеран“, на предметима: Увод у програмирање, Програмирање 1, Програмирање 2, Структуре података и алгоритми и *XML* технологије. У току рада на факултету активно је учествовала у великом броју европских пројеката на којима се бавила развојем мобилних, веб и десктоп апликација.

Докторске студије на Факултету организационих наука уписала је 2015 године, на студијском програму Информациони системи и квантитативни менаџмент, изборно подручје Софтверско инжењерство. Положила је следеће испите:

1. Прикупљање софтверских захтева – оцена 9 (девет)
2. Пројектовање софтвера – одабрана поглавља – оцена 10 (десет)
3. Софтверски процес и одржавање софтвера – одабрана поглавља – оцена 10 (десет)
4. Вештачка интелигенција - оцена 10 (десет)
5. Управљање софтверским пројектима - оцена 10 (десет)
6. Квалитет софтвера - оцена 10 (десет)
7. Алати и методе софтверског инжењерства - оцена 10 (десет)
8. Тестирање софтвера- оцена 10 (десет)
9. Конструкција софтвера - оцена 9 (девет)

Кандидаткиња је дана 26.09.2017. одбранила приступни рад за израду дисертације под називом „Софтверски алат за испитивање алгоритама структурне регресије базиране на *GCRF* моделу“.

Списак објављених радова:

1. Радови објављени у научним часописима међународног значаја (M21)
 - 1.1 **Vujičić T.**, Glass J., Zhou F., Obradović Z.: Gaussian Conditional Random Fields Extended for Directed Graphs, Machine Learning Journal, Springer, 2017, ISSN: 0885-6125 (Print), 1573-0565 (Online), DOI: 10.1007/s10994-016-5611-7
2. Радови саопштени на скупу међународног значаја (M33)
 - 2.1 Šćepanović S., **Vujičić T.**, Radunović P., Zupanc K., Bosnić Z.: Geostep – platform for mobile game-based learning, International Electrotechnical and Computer Science IEEE Conference ERK, Portorož, Slovenia, 2015, ISSN: 1581-4572
 - 2.2 Šćepanović S., **Vujičić T.**, Matijević T., Radunović P.: Game based mobile learning – application development and evaluation, Conference on eLearning, Beograd, Srbija, 2015, ISBN: 978-86-89755-07-7
 - 2.3 Šćepanović S., **Vujičić T.**, Radunović P., Antović I., Vlajić S.: Challenges in design of games for mobile learning, INTED2016 Proceedings, Valensija, Španija, 2016, ISBN: 978-84-608-5617-7, ISSN: 2340-1079
 - 2.4 **Vujičić T.**, Šćepanović S., Jovanović J.: Requirements Engineering in Culturally and Technologically Diverse Settings, MECO Conference, Bar, Crna Gora, 2016, ISBN: 978-1-5090-2222-9, DOI: 10.1109/MECO.2016.7525693
 - 2.5 **Vujičić T.**, Matijević T., Ševarac Z., Ljucović J., Balota A.: Comparative Analysis of Methods for Determining Number of Hidden Neurons in Artificial Neural Network, CECIIS Conference, Varaždin, Hrvatska, 2016, ISSN: 1847-2001 (Print), ISSN 1848-2295 (Online)
 - 2.6 Matijević T., **Vujičić T.**, Ljucović J., Radunović P., Balota A.: Performance Analysis of Girvan-Newman Algorithm on Different Types of Random Graphs, CECIIS Conference, Varaždin, Hrvatska, 2016, ISSN: 1847-2001 (Print), ISSN 1848-2295 (Online)

- 2.7 Ljucović J., **Vujičić T.**, Matijević T., Tomović S., Šćepanović S.: Comparative analysis of classic clustering algorithms and Girvan-Newman algorithm for finding communities in social networks, ENTRENOVA, Rovinj, Hrvatska, 2016, ISSN: 1849-7950
 - 2.8 Balota A., Grebović S., **Vujičić T.**, Radunović P.: Connection links and transport of atmospheric discharges registered data, MECO Conference, Bar, Crna Gora, 2017, ISBN: 978-1-5090-6741-1, ISSN 2377-5475, DOI: 10.1109/MECO.2017.7977153
 - 2.9 Radunović P., **Vujičić T.**, Balota A.: Web Application for Lightning Activity Monitoring System (LAMS), CISTI Conference, Lisbon, Portugal, 2017, ISBN: 978-989-98434-7-9
 - 2.10 Žarić N., Šćepanović S., **Vujičić T.**, Ljucović J., Davcev D.: The Model for Gamification of E-learning in Higher Education Based on Learning Styles, ICT Innovations, Skoplje, Makedonija, 2017, ISSN 1865-0929, ISBN 978-3-319-67596-1, DOI 10.1007/978-3-319-67597-8
 - 2.11 Šćepanović S., Loshkovska S., Balota A., Radunović P., Knežević I., **Vujičić T.**, Petrušić D., Antović I.: Case Studies of Dynamic SOA Services for The Healthcare Systems in Montenegro and Macedonia, ICT Innovations Web proceedings, Skoplje, Makedonija, 2017, ISSN 1865-0937
3. Радови саопштени на скупу националног значаја (M63)
 - 3.1 **Vujičić T.**, Radunović P., Karanikić M., Knežević I.: Elektronske ulaznice bazirane na QR kodu, Međunarodni naučno-stručni skup StES (Students Encountering Science), Banja Luka, Bosna i Hercegovina, 2011, ISBN: 999388062-0
 - 3.2 **Vujičić T.**, Strujić Dž., Šendelj R.: Prednosti korišćenja Hibernate frameworka u odnosu na JDBC u Java EE aplikacijama, XVII naučno stručni skup Informacione tehnologije, Žabljak, Crna Gora, 2012, ISBN: 978-86-7664-102-4
 - 3.3 Radunović P., **Vujičić T.**, Karanikić M., Knežević I.: Simulacioni model three-state algoritma, XVII naučno stručni skup Informacione tehnologije, Žabljak, Crna Gora, 2012, ISBN: 978-86-7664-102-4
 - 3.4 **Vujičić T.**, Radunović P., Knežević I.: Komparativna analiza NOSQL i SQL baza podataka, na primjeru Datomic-a i MSSQL-a, XIX naučno stručni skup Informacione tehnologije, Žabljak, Crna Gora, 2014, ISBN: 978-86-85775-15-4
 - 3.5 Radunović P., **Vujičić T.**, Knežević I.: Poređenje funkcionalnog i imperativnog pristupa programiranju, XIX naučno stručni skup Informacione tehnologije, Žabljak, Crna Gora, 2014, ISBN: 978-86-85775-15-4
 - 3.6 **Vujičić T.**, Radunović P., Antović I., Vlajić S.: Upotreba neuronskih mreža za predviđanje upotrebljivosti podataka, YUINFO Conference, Kopaonik, Srbija, 2016, ISBN: 978-86-85525-17-9
 - 3.7 Ljucović J., Matijević T., **Vujičić T.**, Tomović S.: Analysis of social network random model and comparison to real collaboration network, ITRO Conference, Zrenjanin, Srbija, 2016, ISBN 978-86-7672-285-3

1.2 Оцена подобности кандидата за рад на предложеној теми

Кандидаткиња Тијана Вујичић је завршила основне и специјалистичке студије на Факултету за информационе технологије Универзитета „Медитеран“ у Подгорици, а мастер студије на Факултету организационих наука, Универзитета у Београду. На докторским студијама Факултета организационих наука положила је свих девет испита предвиђених планом и програмом студија, одбранила је приступни рад и остварила 120 ЕСПБ бодова.

Узимајући у обзир резултате остварене током досадашњег образовања, резултате досадашњег истраживања објављене у признатим међународним часописима и конференцијама, радно искуство кандидаткиње, као и искуство у држању наставе из области из које је предложена тема, сматрамо да кандидаткиња Тијана Вујичић поседује потребно знање и искуство за израду докторске дисертације на тему „Софтверски алат за испитивање алгоритама структурне регресије базиране на *GCRF* моделу“.

2. Предмет и циљ истраживања

Машинско учење је једна од области вештачке интелигенције која има за циљ креирање система који су способни да се прилагоде новим ситуацијама и уче из искуства [1]. Основни облици машинског учења су надгледано (*supervised*) учење и ненадгледано (*unsupervised*) учење [2]. Традиционални модели за надгледано учење су моћан алат за учење нелинеарних мапирања, али су они усредсређени на предвиђање једног излаза, третирају све улазне-излазне парове независно и не могу искористити везе између објеката како би унапредили предвиђање. Као проширење традиционалних модела креирани су модели структурног предвиђања [3], који су дизајнирани да користе везе између објеката при предвиђању излазних вредности. Проблеми који се могу решити коришћењем структурног предвиђања могу се представити као графови, у којима сваки чвор представља објекат који има атрибуте (x) и излазну вредност (y), док везе између чворова зависе од конкретне примене и одређени су унапред, преко доменског знања или претпоставке. На пример, веза између болница може се базирати на сличности њихових специјализација, веза између научних радова може се одредити на основу сличности радова који их цитирају, веза између два лица може представљати јачину њиховог пријатељства, итд. Матрица повезаности графа користи се приликом предвиђања и назива се матрицом сличности (S). Постоји велики број радова у просторној статистици која има за циљ истраживање корелације у структурним подацима [4, 5, 6, 7]. Већина истраживања у овој области односи се на предвиђање дискретних излаза (класификација), док је примена структурних модела за проблеме регресије мање истражена област.

Модел *Gaussian Conditional Random Fields (GCRF)* [8] је један од најчешће коришћених модела структурне регресије који интегрише предикцију традиционалних модела надгледаног учења (неструктурних предиктора) и везу између објеката у циљу тачније предикције. Као неструктурни предиктори могу се користити на пример неуронске мреже [9] или линеарна регресија [10]. Овај модел је први пут примењен у области рачунарске визије (*computer vision*) [11] и од тада је коришћен у различитим областима као што су: метеорологија [12], енергетика [13], здравство [14], препознавање говора [15], али и проширен за различите намене [16, 17, 18, 19]. Главна претпоставка овог модела је да су два објекта која су уско повезана веома слични један другом, па самим

тим и вредности њихових излазних варијабли треба бити сличне. Сличност између објеката у *GCRF* моделу мора бити симетрична, али у великом броју реалних проблема објекти су несиметрично повезани. На пример, јачина пријатељства често није симетрична. У емпиријским студијама од испитаника се обично тражи да наведу своје пријатеље и да одреде колико су с њима блиски, што резултира усмереним графом у ком често постоје једносмерна пријатељства [20, 21]. Још један пример усмереног графа су друштвене мреже *Twitter*, *Instagram* или *GitHub*, у којима веза „*follower-follower*“ не мора бити, и често није, двосмерна.

Циљ докторске дисертације је предлог и евалуација усмереног *GCRF* модела и развој апликације која интегрише различите врсте *GCRF* модела. Евалуација модела укључује тестирање прецизности модела на различитим синтетичким и реалним подацима, анализу перформанси модела и анализу његове конвексности. Апликација треба да омогући тренирање и тестирање *GCRF* модела на различитим сетовима података преко корисничког интерфејса, који треба да буде разумљив и једноставан и да поједностави коришћење *GCRF* модела за кориснике који немају искуства у машинском учењу, али и за истраживаче који желе да упореде своје моделе са *GCRF*-ом. Такође, биће развијена и Јава библиотека која омогућава истраживачима који имају искуства у Јава програмирању да користе *GCRF* модел у свом коду и да једноставно имплементирају његове проширења.

2.1 Списак изабраних референци

- [1] Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development
- [2] Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). Foundations of Machine Learning, The MIT Press
- [3] BakIr, G., Taskar, B., Hofmann, T., Schölkopf, B., Smola, A., & Vishwanathan, S. (2007). Predicting Structured Data, MIT Press
- [3] Anselin, L. (1988). Spatial Econometrics: Methods and Models, Kluwer Academic
- [4] Bo, L., & Sminchisescu, C. (2009). Twin Gaussian Processes for Structured Prediction, International Journal of Computer Vision, no. 1-2, pp. 28-52
- [5] Solberg, A. H. S., Taxt, T., & Jain, A. K. (1996). A Markov random field model for classification of multisource satellite imagery, IEEE Transactions on Geoscience and Remote Sensing, no. 1, pp. 100-113
- [7] Lafferty, J., McCallum, A., & Pereira, F. (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, 18th International Conference on Machine Learning
- [8] Radosavljevic, V., Vucetic, S., & Obradovic, Z. (2010). Continuous conditional random fields for regression in remote sensing, ECAI, pp. 809–814
- [9] Haykin, S. (2009). Neural networks and learning machines, Vol. 3, Upper Saddle River Pearson
- [10] Weisberg, S. (2005). Applied linear regression, Vol. 528, Hoboken: Wiley
- [11] Liu, C., Adelson, E. H., & Freeman, W. T. (2007). Learning Gaussian conditional random fields for low-level vision, CVPR Citeseer

- [12] Radosavljevic, V., Vucetic, S., & Obradovic, Z. (2014). Neural Gaussian conditional random fields, Joint European conference on machine learning and knowledge discovery in databases, Springer
- [13] Wytock, M., & Kolter, J. Z. (2013). Sparse Gaussian conditional random fields: Algorithms, theory, and application to energy forecasting, ICML
- [14] Gligorijevic, D., Stojanovic, J., & Obradovic, Z. (2015). Improving confidence while predicting trends in temporal disease networks, 4th workshop on data mining for medicine and healthcare, SIAM international conference on data mining (SDM)
- [15] Khorram, S., Bahmaninezhad, F., & Sameti, H. (2014). Speech synthesis based on Gaussian conditional random fields, Artificial intelligence and signal processing
- [16] Glass, J., Ghalwash, M., Vukicevic, M., & Obradovic, Z. (2016). Extending the modelling capacity of gaussian conditional random fields while learning faster, AAAI
- [17] Stojanovic, J., Gligorijevic, D., & Obradovic, Z. (2015). Semi-supervised learning for structured regression on partially observed attributed graphs, SIAM International Conference on Data Mining
- [18] Gligorijevic, D., Stojanovic, J., & Obradovic, Z. (2016). Uncertainty propagation in long-term structured regression on evolving networks, AAAI
- [19] Han, C., Zhang, S., Ghalwash, M., Vucetic, S. & Obradovic, Z. (2016). Joint learning of representation and structure for sparse regression on graphs, SIAM International Conference on Data Mining
- [20] Michell, L., & Amos, A. (1997). Girls, pecking order and smoking, Social Science & Medicine, pp. 1861–1869
- [21] Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics, Social Networks, pp. 44–60

3. Полазне хипотезе

На основу анализе доступне литературе и постављеног предмета и циља истраживања могу се дефинисати хипотезе докторске дисертације, које ће у раду бити потврђене или оповргнуте:

- Постојећа проширења *GCRF* модела се не могу применити на усмерене графове.
- Предложено проширење *GCRF* модела за усмерене графове омогућава прецизније предвиђање од стандардног *GCRF* модела и традиционалних неструктурних предиктора.
- Софтверски алат поједностављује тренирање и тестирање *GCRF* модела и његових проширења на различитим сетовима података.
- Софтверски алат који нуди имплементацију стандардног *GCRF* модела и његових проширења могу користити како и експерти у области машинског учења, тако и почетници којима *GCRF* модел може помоћи да дођу до жељених информација.

4. Научне методе истраживања

Да би се успешно реализовала идеја истраживања и потврдиле (или оповргле) постављене хипотезе, у раду ће бити коришћен основни метод истраживања који се

базира на прикупљању, анализи и класификацији постојеће литературе и теоријских и експерименталних резултата.

За анализу области и постојећих *GCRF* модела користиће се општи методи анализе постојећих научних резултата, на основу доступне литературе и објављених радова који приказују моделе и њихову примену у пракси.

За проширење *GCRF* модела ће се предложити математички модел који ће бити имплементиран у Јава програмском језику и тестиран на различитим врстама синтетички генерисаних усмерених графова и усмерених графова из реалних сетова податка. Перформансе и прецизност креираног модела ће бити упоређени са постојећим *GCRF* моделом, као и са традиционалним неструктурним предикторима. Експерименти ће се спроводити на истој хардверској и софтверској инфраструктури, на што већем броју примера, чиме ће се омогућити транспарентност и поновљивост експеримената.

Апликација ће се развијати у складу са савременим методама софтверског инжењерства који се примењују приликом развоја софтвера отвореног кода.

Резултати истраживања ће бити презентовани текстуално и графички, коришћењем слика, табела и дијаграма, док ће се сав код бити отворен и доступан на *GitHub*-у

5. Очекивани научни допринос

Као очекивани резултати овог рада могу се навести:

- Целовит приказ и анализа постојећих *GCRF* модела
- Развој математичког модела за проширење *GCRF* модела за усмерене графове
- Имплементација и евалуација новог *GCRF* модела за усмерене графове
- Развој апликације која интегрише различите врсте *GCRF* модела
- Евалуација апликације са корисницима различитих профила и различитих знања из области машинског учења
- Развој Јава библиотеке за *GCRF* моделе

6. План истраживања и структура рада

У уводном делу дисертације ће бити дефинисани предмет и проблем истраживања, циљеви истраживања, хипотезе и методе истраживања. У другом поглављу ће бити дат преглед релевантних области: машинског учења, структурне регресије, *GCRF* модела и његових проширења. У трећем поглављу ће бити описано проширење *GCRF* модела за усмерене графове: математички модел, имплементација и евалуација, која укључује тестирање прецизности предвиђања на различитим скуповима података, као и анализу перформанси и конвексности модела. У четвртом поглављу ће бити представљен софтверски алат који нуди имплементацију стандардног *GCRF* модела и његових проширења. Биће представљене све функционалности алата, његов дизајн и начин имплементације, као и начин коришћења алата за тренирање различитих модела на различитим сетовима података. Такође, биће представљени и резултати евалуације алата са корисницима различитих профила. У петом поглављу биће описана имплементација и начин коришћења Јава библиотеке која садржи основне класе за *GCRF* концепте и имплементацију неких конкретних модела. Шесто поглавље ће садржати примере примене, у седмом поглављу ће бити приказана анализа предности и

недостатака, као и план развоја, док ће у осмом поглављу бити дата завршна разматрања.

Оквирни садржај докторске дисертације је дат у наставку.

1. Увод
 - 1.1 Проблем, предмет и циљ истраживања
 - 1.2 Полазне хипотезе
 - 1.3 Метод истраживања
 - 1.4 Преглед садржаја рада
2. Преглед области
 - 2.1 Машинско учење
 - 2.2 Структурна регресија
 - 2.3 *GCRF* модел
 - 2.4 Проширења *GCRF* модела
3. Усмерени *GCRF*
 - 3.1 Анализа проблема
 - 3.2 Математички модел
 - 3.3 Имплементација модела
 - 3.4 Тестирање модела
 - 3.4.1 Тестирање модела на синтетичким подацима
 - 3.4.2 Тестирање модела на реалним подацима
 - 3.4.3 Перформансе модела
 - 3.4.4 Конвексност модела
4. Софтверски алат
 - 4.1. Функционални опис
 - 4.1.1 Инсталација
 - 4.1.2 Конфигурација
 - 4.1.3 Сетови података
 - 4.1.4 Тренинг и тестирање модела
 - 4.1.5 Документација
 - 4.2 Дизајн и имплементација
 - 4.2.1 Архитектура решења
 - 4.2.2 Имплементација предиктора
 - 4.2.3 Имплементација модела
 - 4.2.4. Имплементација алгоритма за учење
 - 4.2.5 Генерисање синтетичких података
 - 4.2.6 Повезивање са MatLab-ом
 - 4.3 Евалуација софтверског алата
 - 4.3.1 Основне информације о корисницима
 - 4.3.2 Једноставност коришћења
 - 4.3.3 Терминологија
 - 4.3.4 Употребљивост система
 - 4.3.5 Анализа резултата
5. Јава библиотека *GCRFs*
 - 5.1 Имплементација
 - 5.2 Начин употребе
6. Примери примене

7. Дискусија

7. Закључак и предлог

Из изложеног се може закључити да кандидаткиња Тијана Вујичић поседује неопходна знања за успешан рад на докторској дисертацији и испуњава све услове, предвиђене Законом о високом образовању и Статутом Факултета, за одобрење израде докторске дисертације под насловом „Софтверски алат за испитивање алгоритама структурне регресије базиране на *GCRF* моделу“. Предложена тема спада у научна подручја која се развијају на Факултету организационих наука. Тема докторске дисертације по предмету истраживања, садржају и очекиваним доприносима представља значајно подручје истраживања, како са теоријског, тако и са аспекта примене у пракси, у области Техничких наука, уже научне области Софтверско инжењерство.

Комисија је констатовала да кандидаткиња Тијана Вујичић испуњава формалне услове и поседује све потребне квалификације за одобравање израде докторске дисертације на предложеној теми.

На основу изнетог, Комисија за оцену испуњености услова кандидата и научне заснованости теме докторске дисертације предлаже Наставно-научном већу Факултета организационих наука да кандидаткињи Тијани Вујичић одобри израду дисертације под називом:

**„Софтверски алат за испитивање алгоритама структурне
регресије базиране на *GCRF* моделу“**

и да се за ментора именује др Владан Девеџић, редовни професор Факултета организационих наука.

ЧЛАНОВИ КОМИСИЈЕ

Проф. др Владан Девеџић, редовни професор - ментор
Универзитета у Београду, Факултет организационих наука

др Вељко Јеремић, доцент
Универзитета у Београду, Факултет организационих наука

Проф. др Иван Обрадовић, редовни професор у пензији
Универзитета у Београду, Рударско геолошки факултет