

Изборном већу
Математичког факултета
Универзитета у Београду

На 110. седници Изборног већа Математичког факултета Универзитета у Београду одржаној 29.09.2023. одређени смо за чланове комисије за писање извештаја о кандидатима пријављеним на конкурс за избор једног ванредног професора за ужу научну област Рачунарство и информатика. На конкурс који је објављен 11.10.2023. године у издању листа *Послови Националне службе за запошљавање*, у предвиђеном року пријавио се један кандидат: др Јелена Граовац. Комисија је разматрала пријаву кандидата пријављеног на конкурс, и након разматрања приложеног материјала подносимо следећи извештај као и завршни предлог.

Извештај

о пријављеном учеснику на конкурс за избор једног наставника у звање ванредног професора за ужу научну област Рачунарство и информатика

1 др Јелена Граовац

1.1 Биографија

Јелена Граовац (рођена Томашевић) рођена је 14.12.1979. године у Приштини. Основну школу и гимназију завршила је у Приштини са просечном оценом 5.00. Студије на Математичком факултету у Београду, на смеру Рачунарство и информатика уписала је 1999. године и завршила је 2004. године са просеком оцена 9.62. Магистарске студије је уписала на истом факултету 2005. године. Магистрирала је 2008. године, одбравивши магистарски рад под насловом *XML базе података у управљању лексичким ресурсима*, а докторирала је 2014. године у области рачунарских наука на Математичком факултету Универзитета у Београду, одбравивши докторску тезу под насловом *Прилог методама класификације текста: математички модели и примене*. Била је добитник стипендије краљевине Норвешке за 500 најбољих студената у Србији 2002. године, стипендије Фондације за развој научног и уметничког подмлатка, Министарства просвете Републике Србије (2002-2004. године), као и похвале Математичког факултета за постигнуте изузетне успехе током студија 2004. године.

Запослена је на Катедри за рачунарство и информатику Математичког факултета Универзитета у Београду у звању асистента приправника (од 2004. године), сарадника у настави (од 2007. године), асистента (од 2008. године) и доцента (од 2015. године). Током школске 2016/2017. године била је координатор за односе са привредом Математичког факултета.

1.2 Наставна делатност

Јелена Граовац је успешно држала предавања и вежбе из следећих предмета: Пројектовање база података, Информациони системи, Веб програмирање, Увод у организацију и архитектуру рачунара 1, Увод у организацију рачунара, Техничко и научно писање, Основи рачунарских система, Објектно оријентисано програмирање, Програмирање 1, Програмирање 2, Основи програмирања и Интелигентно претраживање (на докторским академским студијама "Интеллигентни системи" Универзитета у Београду). Петогодишњи просек оцена на студентским анкетама је 4.78 (изузимајући академске године 2022/2023 и 2017/2018, због одсуства).

Јелена Граовац је коаутор једне збирке задатака и три скрипте за курсеве студија рачунарства и информатике:

- Milena Vujošević Janičić, Jelena Graovac, Nina Radojčić, Ana Spasić, Mirko Spasić, Anđelka Zečević.
Programiranje 2 - Zbirka zadataka sa rešenjima (u programskom jeziku C).
Izdavač: Matematički fakultet, 2016.
ISBN: 978-86-7589-107-9
- Jelena Graovac, Milena Vujošević Janičić,
Stručna i naučna komunikacija (skripta)
<http://www.itkomunikacija.matf.bg.ac.rs/skripta/snki.pdf>
- Jelena Graovac,
Projektovanje baza podataka (skripta)
http://poincare.matf.bg.ac.rs/~jgraovac/courses/projbp/2016_2017/projbp_skripta.pdf
- Milena Vujošević Janičić, Jelena Graovac,
Osnovi računarskih sistema - C++ (skripta)
<https://www.matf.bg.ac.rs/~jelena.graovac/courses/ors-zadaci/KompletniMaterijaliORS.zip>

1.3 Научна делатност

Области научног интересовања Јелене Граовац су обрада природних језика, машинско учење и вештачка интелигенција. До сада је објавила пет радова на SCI листи од којих је један самосталан, два рада у међународним часописима и три рада у националним часописима. Имала је више саопштења на међународним и домаћим конференцијама. Од 2018. године до данас, члан је уредничког одбора међународног часописа *Intelligent Data Analysis*. Јелена Граовац је учествовала у пројектима 1858, 144030, 174021 и III47003 Министарства просвете, науке и технолошког развоја. До сада је руководила израдом четири мастер рада, била коментор једне докторске дисертација, и члан комисије за оцену и одбрану једне докторске дисертације и 24 мастер радова. Тренутно руководи израдом једне докторске дисертације. До сада је одржала једно предавање по позиву и била предавач на летњој школи са темом везаном за технике класификације текста. Суоснивач је Друштва за језичке ресурсе и технологије (JERTEX) у чијем је раду активно учествовала.

1.3.1 Радови у међународним часописима са SCI листе

1. M.Šošić, J. Graovac: **Effective methods for email classification: Is it a business or personal email?**. *Computer Science and Information Systems*, vol. 19, no. 3, pp. 1155-1175, 2022. (M23)
2. J. Graovac, M. Mladenović, I. Tanasijević: **NgramSPD: Exploring Optimal N-gram Model for Sentiment Polarity Detection in Different Languages**. *Intelligent Data Analysis*, vol. 23, no. 2, pp. 279-296, 2019. (M23)
3. J. Graovac, J. Kovačević, G. Pavlović-Lažetić: **Hierarchical vs. flat n-gram-based text categorization: can we do better?**. *Computer Science and Information Systems*, vol. 14, no. 1, pp. 103-121, 2017. (M23)

4. J. Graovac: A variant of n-gram based language-independent text categorization. *Intelligent Data Analysis*, vol. 18, no. 4, pages 677–695, 2014. (M23)
5. M. Vujošević Janičić, J. Tomašević, P. Janičić: **Random k-GD-SAT Model and its Phase Transition**. *Journal of Universal Computer Science*, vol. 13, issue 4, pp 572–591, 2007. (M23)

1.3.2 Поглавља у књигама

1. P. Quaresma, P. Janičić, J. Tomašević, M. Vujošević-Jančić, D. Tošić: XML-based Format for Descriptions of Geometrical Constructions and Geometrical Proofs. *Communicating Mathematics in Digital Era, CMDE 2006*, Aveiro, Portugal, August 15-18, 2006.

1.3.3 Радови у научним часописима који нису на SCI листи

1. J. Kovačević, J. Graovac: Prospective automated hierarchical classification of digitized documents, *Review of the National Center for Digitization*, vol. 29, pp. 42–51, 2016. (M53)
2. J. Kovačević, J. Graovac: Application of a Structural Support Vector Machine method to N-gram based text classification in Serbian. *INFOtheca - Journal of Information and Library Science*, vol. 16, no. 1, 2015. (M53)
3. J. Graovac, J. Kovačević, G. Pavlović-Lažetić: Language Independent n-Gram-Based Text Categorization with Weighting Factors: A Case Study. *JIDM - Journal of Information and Data Management*, vol. 6, no. 1, pp 4–17, 2015. (M24)
4. J. Graovac: Wordnet-based text categorization technique, *INFOtheca - Journal of Information and Library Science*, vol. 14, no. 2, pp. 2–17, 2014. (M53)
5. M. Živković, S. Malkov, S. Zarić, M. Vujošević Janičić, J. Tomašević, G. Predović, N. Blažić, M. V. Beljanski: Statistical Dependence of Protein Secondary Structure on Amino Acid Bigrams, *Chemical Industry and Chemical Engineering Quarterly*, 12(1), 82, pp. 82–85, 2006. (M24)

1.3.4 Предавања по позиву

1. J. Graovac: Machine Learning Approach for Sentiment Polarity Detection in Different Languages, *International conference NMLSP 2022: NUMERICAL METHODS FOR LARGE SCALE PROBLEMS*, Belgrade, Serbia, 2022.

1.3.5 Радови у зборницима радова са међународних научних скупова објављени у целини

1. J. Graovac, I. Tomašević, and G. Pavlović-Lažetić: Meet Your Email Sender – Hybrid Approach to Email Signature Extraction. *Asian Conference on Intelligent Information and Da-*

Database Systems. Lecture Note in Artificial Intelligence (LNAI), Springer Nature Switzerland, pp. 547–558, 2022.

2. J. Graovac, M. Radovic, B. Altinel Girgin: ML-SPD: Machine Learning based Sentiment Polarity Detection, *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, pp. 1–7, 2020.
3. J. Graovac, G. Pavlović-Lažetić: Language-Independent Sentiment Polarity Detection in Movie Reviews: A Case Study of English and Spanish, *6th International Conference ICT Innovations 2014*, Web Proceedings, ISSN 1857-7288, pp. 13–22, Ohrid, Macedonia, 2014.
4. J. Graovac: Text categorization using n-gram based language independent technique, *Natural Language Processing for Serbian-Resources and Applications*, In Proceedings of the Conference 35th Anniversary of Computational Linguistics in Serbia, 2014.
5. G. Pavlović-Lažetić, J. Tomašević: Ontology-driven conceptual document classification, *KDIR 2010 - International Conference on Knowledge Discovery and Information Retrieval*, 25-28 10, Valencia, Spain; In Proceedings of the International Conference on Knowledge Discovery and Information Retrieval, pp. 383–386, 2010.
6. J. Tomašević, G. Pavlović-Lažetić: Productivity of concepts in Serbian Wordnet. *11th International Multiconference INFORMATION SOCIETY - IS 2008*, Volume C: Proceedings of the Sixth Language Technologies Conference, EDs. Tomaz Erjavec, Jerneja Zganec Gros, Oct. 16th - 17th, 2008, Ljubljana, Slovenia, pp. 86–91, 2008.

1.3.6 Саопштења на међународним и националним скуповима

1. J. Graovac, J. Kovačević and G. Pavlović-Lažetić: Machine learning-based approach to help diagnosing Alzheimer's disease through spontaneous speech analysis. *Belgrade BioInformatics Conference 2016*, Book of Abstracts, Belgrade, Serbia, 2016
2. J. Graovac, Serbian text categorization using byte level n-grams. *In Proceedings of CLoBL 2012: Workshop on Computational Linguistics and Natural Language*, 5th Balkan Conference in Informatics, ISBN: 978-86-7031-200-5, pp. 93–97, Novi Sad, Serbia, 2012.
3. саопштење: J. Tomašević, M. Vujošević Janičić: TemidaLib – Multiprecision Arithmetic Library. *Zbornik radova SYMOPIS 2006*, Banja Koviljača, 2006.
4. саопштење: M. Vujošević Janičić, J. Tomašević: Phase Transition In Random SAT Problems, *Zbornik radova SYMOPIS 2006*, Banja Koviljača, 2006.
5. постер презентација: M. Vujošević Janičić, J. Tomašević, M. Živković, S. Malkov, G. Predović, N. Blažić, M. V. Beljanski, S. Zarić: Distribution of distances of connected amino acid pairs in proteins. *1st South East European Congress of Chemical Engineering*, Faculty of Technology and Metallurgy, Belgrade, Serbia and Montenegro, 2005.

1.4 Приказ научних радова објављених у часописима на SCI листи

У раду [1] предложен је нов метод за класификацију електронских порука у две категорије: пословне и личне поруке. Примењене су различите методе дубоког учења (енг. Bidirectional Long-Short Term Memory (BiLSTM) и Attention-based BiLSTM (BiLSTM-Att)) као и

алгоритми традиционалног машинског учења (метод подржавајућих вектора оптимизован стохастичним градијентним спустом (енг. Stochastic Gradient Descent) и метод изузетно случајних стабала (енг. Extremely Randomized Trees)). Посебан допринос рада огледа се у екстракцији различитих лексичких, конверзационих, изражајних, емотивних и моралних карактеристика које су се показале веома корисним у разликовању личних од пословних електронских порука. Експерименти су извршени на јавно доступном Епрон корпусу на којем смо постигли резултате који су најбољи у области.

У раду [2] предложен је нов метод за сентимент класификацију текста према поларитету исказаног мишљења које у себи носи. Примењене су различите методе машинског учења (Максимална ентропија, Метод подржавајућих вектора и N најближих суседа). Главни акценат у раду стављен је на испитивање различитих начина репрезентације текста и њиховог утицаја на резултате класификације. Развијени су различити n -грам модели на нивоу речи, карактера и бајта који су у комбинацији са различитим методама машинског учења примењени на најпопуларнијим јавно доступним скуповима филмских рецензија на 7 језика: енглеском, шпанском, арапском, француском, турском, чешком и српском језику. Истраживања су показала да n -грами на нивоу бајта и карактера дају боље резултате од n -грама на нивоу речи, а у случају арапског, француског, турског, чешког и српског језика дају боље резултате од свих осталих до сада објављених метода примењених над овим скуповима.

У раду [3] предложене су две методе за хијерархијску класификацију текста засноване на машинском учењу и n -грамима бајтова. Методе су примењене на најпопуларнијим јавно доступним скуповима текстова на енглеском и кинеском језику. Циљ је био упоредити добијене резултате са резултатима уобичајене, тзв. равне класификације над истим скуповима. Формалном евалуацијом добијених резултата утврђено је да хијерархијска класификација не доноси велико побољшање резултата у односу на равну класификацију, када се обе примене над истим скуповима података.

У раду [4] је предложена нова n -грамска метода репрезентације текста у циљу његове тематске класификације коришћењем метода машинског учења. Развијен је нови метод заснован на n -грамима бајтова и примењен је на јавно доступним скуповима на енглеском, кинеском и српском језику, чиме је илустрована његова потпуна језичка независност. Резултати који су добијени су надманили до тада објављене методе класификације текста примењене над истим скуповима података.

У раду [5] је представљен нов тип SAT проблема, назван k -gd-sat, који уопштава k -sat и gd-sat проблеме. У њему дужина клаузе има геометријску расподелу контролисану вероватносним параметром p ; ако је $p = 1$, k -gd-sat проблем се своди на k -sat проблем. Показано је да постоји фазна промена између задовољности и незадовољности за случајно генерисане инстанце овог проблема. У раду су приказани резултати теоријске анализе и експеримената који су сугерисали да постоји линеарни однос по параметру $1/p$ између тачака укрштања за различите параметре k -gd-sat проблема. Разматран је и однос између тачака укрштања за k -sat и k -gd-sat проблеме.

1.5 Приказ научних радова објављених у часописима који нису на SCI листи

У раду [1] је представљена метода хијерархијске класификације докумената НЦД (енг. National Center for Digitization) дигиталне библиотеке заснована на методи структуралних подржавајућих вектора. Приказана је развијена метода, предложена су два типа хијерархија

класа НЦД библиотеке на основу њиховог садржаја и дефинисан је протокол за примену ове методе на дигитализоване документе.

У раду [2] су приказани резултати класификације хијерархијски организованог EBART корпуса на српском језику. Коришћене су две технике засноване на методи структуралних подржавајућих вектора: вишекласна равна и хијерархијска класификација. Резултати су показали да хијерархијска класификација даје боље резултате за специфичне класе са малим бројем текстова, док су добијени слични резултати када су методе примењене на цео корпус. У овом раду су приказани резултати класификације хијерархијски организованог EBART корпуса на српском језику. Коришћене су две технике засноване на методи структуралних подржавајућих вектора: вишекласна равна и хијерархијска класификација. Резултати су показали да хијерархијска класификација даје боље резултате за специфичне класе са малим бројем текстова, док су добијени слични резултати када су методе примењене на цео корпус.

У раду [3] је представљена језички независна техника класификације текстова заснована на профилима n-грама бајтова са придруженим тежинским факторима и поређењу ових профила. Техника је примењена на 2 веома различита, широко распрострањена светска језика: енглеском и арапском, чиме је демонстрирана језичка независност методе. Резултати који су постигнути за јавно доступне корпусе су показали да је ова техника надмашила све доступне резултате на арапском језику и постигла резултате упоредиве са најбољим резултатима на енглеском језику.

У раду [4] је приказана техника категоризације текстова заснована српском Wordnet-у.

У раду [5] је приказана статистичка зависност секундарне структуре изабраних протеина из Protein Data Bank базе на основу фреквенције биграма аминокиселина.

2 Закључно мишљење и предлог комисије

Кандидат др Јелена Граовац (рођена Томановић) завршила је основне студије на Математичком факултету на смеру Рачунарство и информатика са просечном оценом 9.62. Магистарске студије из области рачунарства је завршила 2008. године са просечном оценом 10. Докторат из области рачунарства је одбранила 2014. године на истом факултету са докторском дисертацијом *Прилог методама класификације текста: математички модели и примене*. Јелена Граовац је у радном односу на Математичком факултету од 2004. године и држала је вежбе и предавања из великог броја предмета на основним и мастер студијама. Петогодишњи просек оцена на студентским анкетама др Јелене Граовац је 4.78 (од 5). Руководила је израдом четири мастер тезе и била коментор једног докторског рада. Коаутор је 1 збирке задатака и 4 скрипте за студије рачунарства и информатике. Укупно је објавила 5 научних радова у часописима на SCI листи (3 рада у периоду од првог избора за доцента). Одржала је једно предавање по позиву и имала већи број излагања на међународним и националним научним скуповима (од првог избора за доцента излагала је 4 пута на међународним скуповима). Др Јелена Граовац је учествовала на више националних научних пројеката.

На основу свега изложеног, комисија сматра да кандидат др Јелена Граовац испуњава све формалне и суштинеке услове конкурса и **предлаже Изборном већу Математичког факултета у Београду да др Јелену Граовац изабере у звање ванредног професора за ужу научну област Рачунарство и информатика.**

У Београду,
18.11.2023. године

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

(др Предраг Јашичић, ред. проф.)

(др Ненад Митић, ред. проф.)

(др Рашка Сташковић, ван. проф. Рударско-геолошки факултет)