

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Подобност теме и кандидата Вука Батановића, мастер инж. електротехнике и рачунарства, за израду докторске дисертације

Одлуком бр. 5045/11-1 од 28.04.2016. године, именовани смо за чланове Комисије за оцену подобности теме и кандидата Вука Батановића, мастер инж. електротехнике и рачунарства, за израду докторске дисертације и научне заснованости теме „Методологија решавања семантичких проблема у обради кратких текстова написаних на природним језицима са ограниченим ресурсима“.

На основу материјала приложеног уз Захтев кандидата, Комисија подноси следећи

ИЗВЕШТАЈ

1. Подаци о кандидату

1.1. Биографски подаци

Вук Батановић је рођен 02.06.1987. године у Београду. Трећу београдску гимназију је завршио као носилац Вукове дипломе 2006. године, када је и уписао Електротехнички факултет у Београду. Дипломирао је на модулу за рачунарску технику и информатику у октобру 2010. са просечном оценом 9,56 док је дипломски рад добио оцену 10. Мастер студије на модулу за рачунарску технику и информатику Електротехничког факултета је уписао истог месеца. Студије је завршио у октобру 2011. са просечном оценом 10 а завршни мастер рад је одбранио са истом оценом. Крајем децембра 2011. године уписао је докторске студије на Електротехничком факултету, модул Софтверско инжењерство. Положио је све испите на докторским студијама са просечном оценом 10. Предлог теме докторске дисертације кандидат је поднео 04.04.2016. године. Научна интересовања су му фокусирана на решавање проблема из области обраде природних језика (енгл. Natural Language Processing), пре свега применом метода машинског учења. Нарочито га интересују специфичности обраде кратких текстова и креирање решења примењивих и у језицима са ограниченим језичким ресурсима.

1.2. Сечено научноистраживачко искуство

Кандидат је у досадашњем току докторских студија испунио све обавезе и положио све испите предвиђене студијским планом и програмом, остваривши укупно 120 ЕСПБ и општи успех 10,00 (десет и 00/100). Испити које је кандидат положио у оквиру докторских студија су:

Назив предмета	Оцена	ЕСПБ
Сигурност и заштита рачунарских система	10 (десет)	9
Принципи програмских језика	10 (десет)	9
Интернет програмирање	10 (десет)	9
Вештачка интелигенција	10 (десет)	9
Вештачка интелигенција и експертски системи	10 (десет)	9
Увод у научни рад	10 (десет)	6
Софтверска техника	10 (десет)	9
Медицинска информатика	10 (десет)	9
Виртуелна стварност	10 (десет)	9
Неуралне мреже	10 (десет)	9

Поред тога, кандидат је одрадио следеће студијске обавезе:

Назив обавезе	ЕСПБ
Научно стручни рад	3
Студијски истраживачки рад I	15
Студијски истраживачки рад II	15

Кандидат је коаутор шест научних радова у вези са темом дисертације, од чега су два рада категорије M20, један рад категорије M30, један рад категорије M50 и два рада категорије M60, као и два техничка решења (категорија M80):

Категорија M20 – Радови објављени у научним часописима међународног значаја

- Категорија M21

1. Furlan B., Batanović V., Nikolić B.: *Semantic similarity of short texts in languages with a deficient natural language processing support*, Decision Support Systems, vol. 55, no. 3, pp. 710–719, 2013. DOI: 10.1016/j.dss.2013.02.002, ISSN: 0167-9236, IF (2013) = 2.036

- Категорија M23

1. Batanović V., Bojić D.: *Using Part-of-Speech Tags as Deep-Syntax Indicators in Determining Short-Text Semantic Similarity*, Computer Science and Information Systems, vol. 12, no. 1, pp. 1–31, 2015. DOI: 10.2298/CSIS131127082B, ISSN: 1820-0214 (Print), 2406-1018 (Online), IF (2015) = 0.623

Категорија M30 – Зборници међународних научних скупова

- Категорија M33

1. Batanović V., Nikolić B., Milosavljević M.: *Reliable Baselines for Sentiment Analysis in Resource-Limited Languages: The Serbian Movie Review Dataset*, in Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016), pp. 2688-2696, European Language Resources Association (ELRA), Portorož, Slovenia, 2016.

Категорија M50 – Часописи националног значаја

- Категорија M52

1. Batanović V., Bojić D.: *Evaluation and Classification of Syntax Usage in Determining Short-Text Semantic Similarity*, Telfor Journal, vol. 6, no. 1, pp. 64–68, 2014. DOI: 10.5937/telfor1401064B, ISSN: 1821-3251 (Print), 2334-9905 (Online)

Категорија М60 – Зборници скупова националног значаја

- Категорија М63

1. Batanović V., Bojić D.: *Evaluacija i klasifikacija korišćenja sintaksnih informacija u određivanju semantičke sličnosti kratkih tekstova*, in Proceedings of the 21st Telecommunications Forum (TELFOR 2013), pp. 821–824, IEEE, Belgrade, Serbia, 2013. DOI: 10.1109/TELFOR.2013.6716356, ISBN: 978-1-4799-1419-7 (Print), 978-1-4799-1420-3 (Online)
2. Batanović V., Furlan B., Nikolić B.: *Softverski sistem za određivanje semantičke sličnosti kratkih tekstova na srpskom jeziku*, in Proceedings of the 19th Telecommunications Forum (TELFOR 2011), pp. 1249–1252, IEEE, Belgrade, Serbia, 2011. DOI: 10.1109/TELFOR.2011.6143778, ISBN: 978-1-4577-1499-3 (Print), 978-1-4577-1500-6 (Online)

Категорија М80 – Техничка и развојна решења

- Категорија М81

1. Batanović V., Furlan B., Nikolić B.: *Softverski sistem za određivanje semantičke sličnosti kratkog teksta napisanog na srpskom jeziku*, Elektrotehnički fakultet u Beogradu, 2013.

- Категорија М86

1. Batanović V., Furlan B., Nikolić B.: *Korpus parafraza srpskog jezika*, Elektrotehnički fakultet u Beogradu, 2013.

Кандидат је успешно одбранио предложену тему докторске дисертације на јавној усменој одбрани одржаној 17.05.2016. године пред комисијом за оцену услова и прихватање теме докторске дисертације у саставу:

1. др Драган Бојић, ванредни професор, Универзитет у Београду – Електротехнички факултет
2. др Зоран Шеварац, доцент, Универзитет у Београду – Факултет организационих наука
3. др Горан Квашчев, доцент, Универзитет у Београду – Електротехнички факултет

1.3. Оцена подобности кандидата за рад на предложеној теми

На основу претходног приказа резултата досадашњег научног рада кандидата може се видети да је он већ имао успеха у развијању решења и објављивању резултата научних доприноса из области рачунарске обраде природних језика, са акцентом на проблеме одређивања семантичке сличности и анализе сентимента текстова, који представљају окосницу планиране докторске дисертације. Стога Комисија процењује да је кандидат доказао способност рада на предложеној теми докторске дисертације.

2. Предмет и циљ истраживања

Под семантичке проблеме у обради природних језика може се подвести већи број задатака који подразумевају или за циљ имају правилно разумевање значења текстова. Предмет истраживања у овој дисертацији представљају два проблема у обради семантике кратких текстова – анализа сентимента текстова и одређивање семантичке сличности два текста. Под појмом „кратких текстова“ се у истраживању подразумевају текстови дужине од неколико речи до пар реченица. Ова два проблема су сродна не само због тога што се оба односе на значење текстова већ и по томе што се оба могу формализовати на сличне начине – било као задаци бинарне класификације (на позитивне/негативне текстове, односно парафразе/не-парафразе) било у облику вишекласне класификације или регресије (где појединачне класе или нумеричке вредности означавају тачке на скали од позитивног до негативног

сентимента, односно од семантички идентичног до семантички различитог садржаја). Осим тога, резултати новијих истраживања у обради енглеског језика показују да је при раду са кратким текстовима у оба проблема могуће избећи потпуну синтаксну анализу, па самим тим и потребу за синтаксним алатима као што су парсери. Другим речима, може се директно прећи са нивоа лексичке семантике на ниво семантике целог текста, и то често без губитака у тачностима алгоритама. Коначно, оба проблема су међу најзаступљенијим темама актуелних истраживања у обради природних језика.

Статистички приступи који користе технике машинског учења су данас доминантни у обради природних језика. Ипак, решења приказана у литератури се најчешће развијају и евалуирају само на енглеском језику за који је, захваљујући залагању истраживача широм света, креиран огроман број различитих језичких ресурса и алата. Стога је количина података доступних за изградњу модела и постојање неопходних језичких алата ретко када проблем при обради текстова на енглеском језику. Насупрот томе, у већини слабије заступљених језика, укључујући српски, доступност ресурса представља једну од кључних тешкоћа. У многим доменима обраде језика је тешко пронаћи језичке алате, а ресурси, ако уопште постоје, најчешће су скромног обима, што негативно утиче на квалитет креираних модела. Ова појава ефективно спречава примену комплексних модела при раду са слабије заступљеним језицима јер ти модели по природи захтевају веће количине података и/или различите облике обраде текста како би могли бити адекватно примењени. Сви ови проблеми су нарочито евидентни при раду са кратким текстовима – овај вид обраде је приметно тежи од обраде дужих текстуалних докумената због мање количине доступних информација у тексту и већег утицаја синтаксних одлика.

Циљ овог истраживања јесте прикупљање ресурса и алата за српски језик који су неопходни за изградњу и евалуацију модела за анализу сентимента и одређивање семантичке сличности кратких текстова, као и креирање основних референтних решења (енгл. *baselines*). Као основна решења би се користиле модификације и унапређења једноставних метода које су се показале успешним у обради енглеског и које је могуће прилагодити специфичностима српског језика. Поред тога, циљ је и дати предлог нове методологије решавања наведених проблема, уз развој нових и модификовање постојећих комплекснијих алгоритама из ове области, тако да при раду на језицима са ограниченим ресурсима остварују боље перформансе у односу на то када се ова ресурсна ограничења не узимају у обзир.

Проблематика рачунарске анализе значења кратких текстова на српском језику је до сада била запостављена у научним истраживањима, те би остваривање предложених циљева овог рада знатно допринело напретку области обраде текстова на српском. Креирани скупови података би били драгоцен ресурс у развоју будућих приступа и метода јер њихово стварање обично подразумева обимну мануелну анотацију. Сем тога, планирано је да методологија предложена у овој дисертацији буде лако примењива и на остале језике са слабо развијеним језичким ресурсима и алатима. Методе анализе сентимента текстова су директно примењиве у великом броју реалних ситуација, што их чини од великог практичног значаја. С друге стране, методе одређивања семантичке сличности могу бити од користи било саме за себе било као незаобилазни саставни део у већем броју других задатака из обраде природних језика, као што су сумаризација текстова, одговарање на питања, дохватање информација, итд. Важност анализе сентимента и одређивања семантичке сличности је посебно наглашена у домену обраде кратких текстова због њихове знатне заступљености на интернету у форми порука и коментара на друштвеним мрежама, описа разноврсних производа и услуга, сажећака вести, итд.

2.1. Релевантни библиографски извори за предложено истраживање

У наставку је дат избор значајнијих научних радова из огромног скупа релевантних радова који се баве проблемима анализе сентимента и одређивања семантичке сличности текстова, као и избор радова који разматрају технике морфолошке нормализације српског језика.

1. Pang, B., Lee, L. & Vaithyanathan, S.: *Thumbs up? Sentiment Classification using Machine Learning Techniques*, in Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), pp. 79–86, Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, 2002.
2. Pang, B. & Lee, L.: *A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts*, in Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004), Article No. 271, Association for Computational Linguistics, Morristown, New Jersey, USA, 2004.
3. Pang, B. & Lee, L.: *Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales*, in Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005), pp. 115–124, Association for Computational Linguistics, Ann Arbor, Michigan, USA, 2005.
4. Koppel, M. & Schler, J.: *The Importance of Neutral Examples for Learning Sentiment*, Computational Intelligence, vol. 22, no. 2, pp. 100–109, 2006.
5. Wang, S. & Manning, C. D.: *Baselines and Bigrams: Simple, Good Sentiment and Topic Classification*, in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012), pp. 90–94, Association for Computational Linguistics, Jeju Island, South Korea, 2012.
6. Socher, R. et al.: *Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank*, in Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013), pp. 1631–1642, Association for Computational Linguistics, Seattle, Washington, USA, 2013.
7. Le, Q. & Mikolov, T.: *Distributed Representations of Sentences and Documents*, in Proceedings of the 31st International Conference on Machine Learning (ICML 2014), pp. 1188–1196, Beijing, China, 2014.
8. Mihalcea, R., Corley, C. & Strapparava, C.: *Corpus-based and Knowledge-based Measures of Text Semantic Similarity*, in Proceedings of the 21st National Conference on Artificial Intelligence (AAAI 2006), pp. 775–780, AAAI Press, Boston, Massachusetts, USA, 2006.
9. Islam, A. & Inkpen, D.: *Semantic Text Similarity Using Corpus-Based Word Similarity and String Similarity*, ACM Transactions on Knowledge Discovery from Data, vol. 2, no. 2, Article No. 10, 2008.
10. Oliva, J., Serrano, J. I., del Castillo, M. D. & Iglesias, Á.: *SyMSS: A syntax-based measure for short-text semantic similarity*, Data & Knowledge Engineering, vol. 70, no. 4, pp. 390–405, 2011.
11. Agirre, E., Cer, D., Diab, M. & Gonzalez-Agirre, A.: *SemEval-2012 Task 6: A Pilot on Semantic Textual Similarity*, in Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM), pp. 385–393, Association for Computational Linguistics, Montreal, Canada, 2012.
12. Agirre, E., Cer, D., Diab, M., Gonzalez-Agirre, A. & Guo, W.: **SEM 2013 shared task: Semantic Textual Similarity*, in Second Joint Conference on Lexical and Computational Semantics (*SEM), pp. 32–43, Association for Computational Linguistics, Atlanta, Georgia, USA, 2013.
13. Agirre, E. et al.: *SemEval-2014 Task 10: Multilingual Semantic Textual Similarity*, in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 81–91, Association for Computational Linguistics, Dublin, Ireland, 2014.
14. Agirre, E. et al.: *SemEval-2015 Task 2: Semantic Textual Similarity, English, Spanish and Pilot on Interpretability*, in Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp. 252–263, Association for Computational Linguistics, Denver, Colorado, USA, 2015.
15. Kešelj, V. & Šipka, D.: *Pristup izgradnji stemera i lematizora za jezike s bogatom fleksijom i oskudnim resursima zasnovan na obuhvatanju sufiksa*. INFOTEKA – časopis za bibliotekarstvo i informatiku, vol. 9, no. 1-2, pp. 21–31, 2008.

16. Milošević, N.: *Stemmer for Serbian language*, arXiv preprint arXiv:1209.4471, 2012.
17. Ljubešić, N., Boras, D. & Kubelka, O.: *Retrieving Information in Croatian: Building a Simple and Efficient Rule-Based Stemmer*, in INFUTURE2007: Digital Information and Heritage, pp. 313–320, Department for Information Sciences, Faculty of Humanities and Social Sciences, Zagreb, Croatia, 2007.
18. Gesmundo, A. & Samardžić, T.: *Lemmatizing Serbian as Category Tagging with Bidirectional Sequence Classification*, in Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012), pp. 2103–2106, European Language Resources Association (ELRA), Istanbul, Turkey, 2012.
19. Agić, Ž., Ljubešić, N. & Merkler, D.: *Lemmatization and Morphosyntactic Tagging of Croatian and Serbian*, in Proceedings of the 4th Biennial International Workshop on Balto-Slavic Natural Language Processing, pp. 48–57, Association for Computational Linguistics, Sofia, Bulgaria, 2013.
20. Ljubešić, N., Klubička, F., Agić, Ž. & Jazbec, I.-P.: *New Inflectional Lexicons and Training Corpora for Improved Morphosyntactic Annotation of Croatian and Serbian*, in Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016), pp. 4264–4270, European Language Resources Association (ELRA), Portorož, Slovenia, 2016.

3. Полазне хипотезе

Основне полазне хипотезе усвојене у овој дисертацији су:

- Анализа сентимента и одређивање семантичке сличности кратких текстова се могу успешно реализовати и без потпуне синтаксне анализе текста, чак и у језицима са сложенијим синтаксним правилима, као што је српски језик – пошто су парсери за српски језик тек у зачетку и није извесно да ће у догледно време бити развијена довољно квалитетна решења за парсирање текстова на српском, неопходно је да се семантика текстова обради помоћу система који су у стању да успешно раде без увида у комплетну синтаксну структуру текста.
- Проблем додатне проређености података (енгл. data sparseness) који је проузрокован радом у морфолошки богатим језицима као што је српски се може умањити применом алата за стемовање или лематизацију речи – неки облик морфолошке нормализације је неопходан да би алгоритми вишег нивоа били у стању да различите флективне облике једне исте речи третирају као један појам.
- Могуће је прибавити довољне количине неопходних текстуалних ресурса на српском језику за изградњу модела заснованих на машинском учењу – креирање скупова података у формату потребном за обучавање и/или евалуацију модела је далеко теже у слабије заступљеним језицима попут српског, где често не постоје интернет сајтови одговарајућег типа и тематике и довољне величине (попут нпр. сајта www.rottentomatoes.com који у домену филмова представља вредан извор за анализу сентимента кратких текстова на енглеском). Сем тога, недостатак одговарајућих ознака на постојећим сајтовима чине ручну анотацију нужном.
- Комплексни модели машинског учења, за чију изградњу су обично потребне веће количине текстуалних података, могу да буду прилагођени тако да успешно функционишу и у условима ограничене доступности ресурса – под комплексним моделима се подразумевају различите форме дубоких и рекурзивних неуралних мрежа које су недавно довеле до напретка у великом броју проблема из обраде природних језика, али уз неопходност коришћења већих скупова података за изградњу модела.

4. Научне методе истраживања

Научне методе истраживања које ће бити примењене у овој дисертацији су:

- Ручна анотација података – због ограничене доступности потребних ресурса на српском језику и природе самих проблема који се разматрају немогуће је аутоматски прикупити анотирани податке који су потребни за изградњу и/или евалуацију модела, те је неопходно да се прикупљени текстуални подаци мануелно анотирају. Да би такви подаци могли да буду од користи, неопходно је установити систем критеријума оцењивања и доследно га следити. Планирано је да се на крају анотације експериментално утврди квалитет коришћених критеријума помоћу степена корелације у оценама анотатора.
- Процена и поређење постојећих решења семантичких проблема у обради кратких текстова написаних на природним језицима, са акцентом на њихову примењивост на језике са ограниченим ресурсима.
- Софтверска имплементација – предвиђено је да се све методе представљене у овој дисертацији имплементирају у виду софтверских пакета, пре свега у програмским језицима Java и Python. При томе ће, по потреби, бити коришћене и већ постојеће софтверске библиотеке имплементирани у овим или другим програмским језицима.
- Експериментална евалуација – провера квалитета предложених алгоритама и различитих вредности њихових параметара ће се спровести експериментално, коришћењем метрика за мерење перформанси које су у литератури најраширеније за посматрану проблематику.

5. Очекивани научни допринос

Очекује се да ће ова дисертација довести до следећих научних доприноса:

- Стварање референтних скупова података за анализу сентимента и одређивање семантичке сличности кратких текстова на српском језику – тренутно не постоје било какви скупови кратких текстова који се односе на проблематику анализе сентимента на српском, те би креирање референтних анотираних скупова умногоме олакшало даљи развој ове области на српском језику, јер би омогућило упоредну евалуацију нових и старих решења. Што се домена одређивања семантичке сличности кратких текстова тиче, једини постојећи скуп на српском језику јесте скуп парафраза који је кандидат израдио у оквиру свог мастер рада. Планирано је да се тај скуп прошири са нивоа бинарних на ниво грануларнијих оцена степена сличности текстова, чиме би се прошириле могућности за моделовање и евалуацију решења из ове проблематике.
- Идентификација постојећих приступа за решавање семантичких проблема у обради кратких текстова који су погодни за примену на језицима са ограниченим ресурсима.
- Креирање основних (енгл. *baseline*) алгоритама за анализу сентимента и одређивање семантичке сличности кратких текстова, одређивање њихових могућности на морфолошки и синтаксно сложенијим језицима попут српског, и унапређење основних решења коришћењем техника и алата одабраних тако да се уваже ове сложености језика. На овај начин би се у погледу перформанси алгоритама при решавању одабрана два проблема поставиле чврсте референтне тачке за даљи развој на српском језику.
- Компаративна евалуација техника морфолошке нормализације текстова на српском – стемовања и лематизације – и различитих решења развијених за ове потребе. Како до сада није спроведено свеобухватно поређење постојећих имплементација ових техника, њихови ефекти на задацима анализе сентимента и одређивања семантичке сличности би били од користи за утврђивање тога која врста и имплементација нормализације је најбоља и под којим условима.

- Утврђивање метода за унапређење перформанси комплексних алгоритама машинског учења при раду са ограниченим ресурсима и експериментална процена ефективности тако прилагођених алгоритама у односу на основна референтна решења.
- Алгоритме развијене у овој дисертацији ће бити једноставно применити и на друге језике са ограниченим језичким ресурсима, тако да се може очекивати да ће представљена методологија бити корисна и у истраживањима рачунарске обраде других мањих језика.

6. План истраживања и структура рада

Следеће активности су планиране у оквиру предложеног истраживања:

- Проналажење погодних извора сирових текстуалних података за анализу сентимента кратких текстова на српском језику
- Прикупљање текстуалних података за анализу сентимента кратких текстова са одабраних извора, њихово форматирање у одговарајућем облику и чишћење од правописних и других грешака.
- Утврђивање критеријума оцењивања и спровођење ручне анотације над прикупљеним подацима за анализу сентимента кратких текстова.
- Утврђивање критеријума оцењивања и проширивање анотације на постојећем скупу парафраза на српском тако да се постигне већа грануларност оцена.
- Евалуација квалитета ручне анотације за обе врсте скупова података.
- Израда основних (енгл. *baseline*) референтних решења за анализу сентимента и одређивање семантичке сличности кратких текстова на српском језику и њихова евалуација коришћењем креираних скупова података.
- Значајно побољшање перформанси основних решења путем испитивања ефеката различитих варијација и подешавања алгоритама.
- Упоредна евалуација ефеката различитих форми морфолошке нормализације текста и њихових специфичних имплементација.
- Одабир одговарајућих комплексних модела машинског учења и њихова примена и евалуација на посматрана два проблема на српском.
- Прилагођење таквих модела условима рада на језицима са ограниченим ресурсима и поређење перформанси прилагођених комплексних модела са резултатима основних референтних решења.

Целокупно истраживање се може поделити на четири сукцесивне фазе, те ће и структура рада бити дефинисана како овим фазама, тако и природом семантичких проблема које дисертација обрађује. Прва фаза је фаза прикупљања сирових текстуалних података. У другој фази се спроводи анотација података прикупљених у претходној фази, као и евалуација квалитета те анотације. Трећа фаза подразумева изградњу, евалуацију и максимирање перформанси основних референтних решења за посматране семантичке проблеме у обради кратких текстова. Коначно, у четвртој фази се врши одабир и прилагођење одговарајућих комплекснијих модела машинског учења условима рада на језицима са ограниченим ресурсима, као и евалуација тако развијених модела.

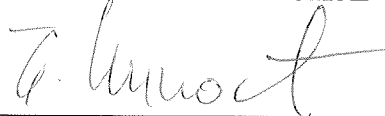
7. Закључак и предлог

На основу претходно изложеног, Комисија је закључила да кандидат Вук Батановић, мастер инж. електротехнике и рачунарства, испуњава све потребне услове за рад на предложеној теми докторске дисертације. Такође, Комисија сматра да је тема предложене докторске дисертације научно заснована и веома актуелна, а делови истраживања су већ верификовани објављивањем резултата истраживања у часописима међународног значаја.

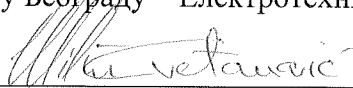
Комисија констатује да тема припада научној области Софтверско инжењерство за коју је Електротехнички факултет Универзитета у Београду матичан. Комисија за менторе докторске дисертације предлаже редовног професора др Бошка Николића и доцента др Милоша Цветановића, запослене на Електротехничком факултету Универзитета у Београду. Комисија предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да кандидату Вуку Батановићу одобри израду докторске дисертације под називом „Методологија решавања семантичких проблема у обради кратких текстова написаних на природним језицима са ограниченим ресурсима“.

У Београду, 13.06.2016.

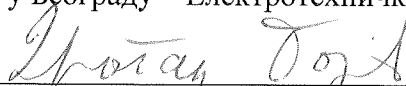
ЧЛАНОВИ КОМИСИЈЕ



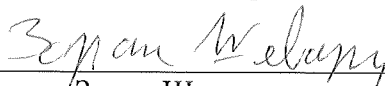
др Бошко Николић, редовни професор
Универзитет у Београду – Електротехнички факултет



др Милош Цветановић, доцент
Универзитет у Београду – Електротехнички факултет



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



др Зоран Шеварац, доцент
Универзитет у Београду – Факултет организационих наука



др Горан Квашчев, доцент
Универзитет у Београду – Електротехнички факултет