

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Бојана Фурлана, дипл. инж.

Одлуком Наставно-научног већа Електротехничког факултета у Београду бр. 5006/07-3 од 5.11.2013. године, именовани смо за чланове Комисије за преглед и оцену докторске дисертације кандидата Бојана Фурлана под насловом

Методологија пројектовања система за интелигентно прослеђивање питања написаних на природном језику

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

Кандидат дипл. инж. Бојан Фурлан уписао је докторске студије на Електротехничком факултету у Београду у школској 2007/2008 години.

Кандидат је пријавио тему за израду докторске дисертације на Електротехничком факултету у Београду 01.03.2013. године. Комисија за студије трећег степена Електротехничког факултета у Београду разматрала је предлог теме за израду докторске дисертације 05.03.2013. године. Наставно – научно веће именovalo је Комисију за оцену услова и прихватање теме докторске дисертације на седници одржаној 12.3.2013. године. у саставу др Бошко Николић, ванредни професор (ментор, Универзитет у Београду – Електротехнички факултет), др Вељко Милутиновић, редовни професор (Универзитет у Београду – Електротехнички факултет), др Ненад Митић, ванредни професор (Универзитет у Београду – Математички факултет). Извештај Комисије за оцену услова и прихватање теме докторске дисертације Наставно – научно веће је прихватило на седници одржаној 14.5.2013. године, а Веће научних области техничких наука Универзитета у Београду дало је сагласност на тему 27.5.2013. године.

Кандидат је предао докторску дисертацију 11.10.2013. године. На седници одржаној 29.10.2013. године Комисија за студије трећег степена дала је сагласност за образовање Комисије за преглед и оцену докторске дисертације. На седници Наставно – научног већа одржаној 05.11.2013. године, именована је Комисија за преглед и оцену докторске дисертације дипл. инж. Бојана Фурлана под насловом „Методологија пројектовања система за интелигентно прослеђивање питања написаних на природном језику“. Чланови комисије су: др Бошко Николић, ванредни професор (ментор, Универзитет у Београду –

Електротехнички факултет), др Вељко Милутиновић, редовни професор (Универзитет у Београду – Електротехнички факултет), др Ненад Митић, ванредни професор (Универзитет у Београду – Математички факултет), др Драган Милићев, ванредни професор (Универзитет у Београду – Електротехнички факултет), др Милош Цветановић, доцент (Универзитет у Београду – Електротехнички факултет).

1.2. Научна област дисертације

Научна област дисертације је Електротехника и рачунарство, а ужа научна област је Рачунарска техника и информатика. За ову ужу научну област матичан је Електротехнички факултет у Београду.

Дисертација је рађена под менторством ванредног професора др Бошка Николића. Ментор испуњава законске услове за ментора, бави се научним радом у ужој области Рачунарска техника и информатика, професионално се бави развојем интелигентних система, а предаје између осталих и предмет Експертски системи на којем се изучавају теме обрађене у дисертацији.

1.3. Биографски подаци о кандидату

Бојан Фурлан је рођен 04.09.1982. године у Панчеву, од оца Дарка и мајке Иванке. Завршио је Средњу Електротехничку Школу „Никола Тесла“ у Београду са одличним успехом. Током школовања био је учесник такмичења у знању, активно тренирао кошарку и бавио се спортом.

Електротехнички факултет Универзитета у Београду уписао је 2001. године, смер Рачунарска техника и информатика. Током студија, био је предавач у Едукационом центру Електротехничког факултета и учесник *Microsoft Student Partner* програма, у оквиру кога је држао курсеве и помагао студентима заинтересованим да овладају Microsoft технологијама.

Након дипломирања и завршетка студија, 2007. године, два пута краће борави на Институту за Информатику, Техничког Универзитета у Минхену, у оквиру DAAD програма Немачке Владе где учествује на истраживачком пројекту SimLab. По повратку уписује докторске студије на Електротехничком факултету, где је у јануару 2008. године примљен у звање асистент у настави за ужу научну област Рачунарска техника и информатика. У периоду од септембра 2009. до јуна 2010. године служи војни рок на Војној Академији, где учествује на развоју информационог система за размену докумената у процесу војног одлучивања.

На Електротехничком факултету у Београду од 2008. године до данас држи вежбе на основним и академским мастер студијама из више предмета. Такође, неколико година је учествовао у формирању, припреми и вођењу екипа које су освојиле бројне награде на такмичењу из програмирања на Електријади.

На докторским студијама положио је све испите са просечном оценом 10. Поред предмета предвиђених наставним планом и програмом похађао је и неколико летњих школа за докторанте из области блиских његовом стручном интересовању.

Аутор је четири рада у међународним часописима, међу којима су и два рада објављена у часописима са *impact* фактором (SCI листа), један из категорије M21 и један из категорије M22. Такође, аутор је једне скрипте са решеним задацима чији је издавач Електротехнички факултет и која се користи у настави на академским мастер студијама. Има више радова који су представљени на међународним и домаћим научним конференцијама, као и техничка решења развијена у оквиру домаћих и међународних пројеката.

Тренутно је рецензент у међународним часописима *Information Sciences* и *Decision Support Systems*, као и домаће конференције ТЕЛФОР. Добитник је награде за најбољи рад конференције IEEE Intelligent Systems IS'12, као и неколико стипендија за краће посете

иностраним научно истраживачким институцијама. Говори енглески и шпански и служи се немачким и руским.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација садржи насловну страну и кратак резиме рада на српском и енглеском језику, садржај, шест поглавља, преглед коришћене литературе, кратку биографију кандидата и прилог. Поглавља су насловљена 1. Увод, 2. Дефиниција проблема, 3. Преглед постојећих и предлог новог решења, 4. Опис и реализација система, 5. Евалуација, 6. Закључак. Дисертација садржи 128 страна (не рачунајући изјаве и прилог), 34 илустрације и 14 табела.

2.2. Кратак приказ појединачних поглавља

Прво поглавље представља увод у дисертацију. Друго поглавље даје опис проблема који се решава, као и осврт на повезаност истраживања из овог домена са осталим блиским областима вештачке интелигенције. На крају овог поглавља дат је оквирни опис фаза СИПП процеса. Треће поглавље садржи детаљан преглед области пројектовања система за интелигентно прослеђивање питања где је посебан акценат стављен на реализацију фаза СИПП процеса. Такође, у овом поглављу је представљена оригинална презентациона парадигма која генерализује суштину свих расположивих СИПП решења из отворене литературе и на основу које је извршена њихова детаљна анализа и евалуација. Као закључак наведен је предлог решења уочених проблема. Четврто поглавље даје приказ дизајна софтверског система који треба да омогући генерализован приступ профилисању корисничког знања, као и аналитички модел који треба да опише карактеристике предложеног решења. У оквиру реализације система приказана је и реализација алгоритама за одређивање семантичке сличности два кратка текста са посебним нагласком на српски језик. Такође, ово поглавље обухвата детаљан приказ реализације софтверског система са становишта имплементације, али и са становишта коришћења, као и преглед најзначајнијих проблема и начина на који су ти проблеми решени. Пето поглавље приказује евалуацију предложеног решења. У оквиру овог поглавља описан је корпус парафраза употребљен за евалуацију алгорита за одређивање семантичке сличности два кратка текста написана на српском језику, као и опис корпуса питања и профила корисника на основу кога је извршена евалуација целокупног система. У шестом поглављу изложен је закључак. На крају су дати неопходни прилози и преглед литературе.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Предмет истраживања у докторској дисертацији је предлог методологије за развој система за интелигентно прослеђивање питања где је посебан акценат стављен на реализацију дефинисаних фаза. У оквиру истраживања дат је предлог и имплементација новог софтверског система који треба да омогући интелигентно прослеђивање питања написаних на природном језику. Систем је предложен на основу оригиналне презентационе парадигме која генерализује суштину свих расположивих СИПП решења. У оквиру предложеног приступа реализован је нови приступ за обраду питања, који омогућава њихову визуализацију. У оквиру рада презентован је нови алгоритам, назван LInSTSS, који приликом одређивања семантичке сличности два кратка текста узима у обзир и

специфичности речи које ови текстови садрже. Резултати добијени евалуацијом рада алгоритма показали су да предложени алгоритам пружа боље резултате у односу на постојећа решења. Такође, предложен је и нови алгоритам за решавање проблема поређења питања и корисничких профила, назван P2Q. Развијени систем је потпуно оригиналан, јер се ниједан од до сада приказаних система не заснива на предложеним алгоритмима и не одликује се свим особинама предложеног система.

3.2. Осврт на референтну и коришћену литературу

Кандидат је детаљно претражио и упознао се са одговарајућом литературом. У дисертацији је прецизно наведено 48 библиографских референци на радове који су у вези са темом дисертације. Литература садржи и најновије радове релевантне за тему дисертације, као и одговарајуће радове чији је аутор или коаутор. Једно поглавље дисертације је посвећено само анализи релевантних приступа, која даје систематизацију и класификацију постојећих решења. Такође, у овом поглављу кандидат уводи оригиналну презентациону парадигму која генерализује суштину свих расположивих решења за интелигентно прослеђивање питања пронађених у отвореној литератури. Искази у том поглављу, као и у осталим деловима дисертације, добро су поткрепљени цитатима одговарајућих радова.

3.3. Опис и адекватност примењених научних метода

Истраживање у оквиру предложене докторске дисертације обухватило је следеће фазе:

1. систематско проучавање домаће и иностране литературе из области дисертације;
2. развој аналитичког СИПП модела;
3. критичку анализу проблема интелигентног прослеђивања питања написаних на природном језику;
4. евалуацију софтверских система за интелигентно прослеђивање питања са становишта три основне фазе које укључује СИПП процес;
5. реализацију прототипа циљног софтверског система и верификацију полазних хипотеза.

Наведени поступци у основи припадају и теоријским и експерименталним истраживањима, и у потпуности одговарају проблему и постављеном циљу дисертације.

3.4. Применљивост остварених резултата

У оквиру дисертације представљена је оригинална методологија за реализацију СИПП система, дефинисане су структура фаза развоја и однос њихових компоненти. Предложени приступ се подједнако може употребити за енглески, српски или неки други језик са врло ограниченим електронским лингвистичким ресурсима. Сваки од реализованих подсистема се може користити као посебан систем, јер омогућава ефикасно решење постављеног проблема. Практична применљивост предложене методологије утврђена је њеном употребом у оквиру развијеног система. Експерименталним путем је утврђено да предложена методологија повећава ефикасност СИПП система. Корист од примене оваквих система укључује смањење

непотребног оптерећења експерата који представљају вредан ресурс и повећање квалитета услуга институције (универзитет, влада, предузеће), имајући у виду задовољство корисника с обзиром да су њихова питања прослеђена релевантним особама.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат је у изради дисертације показао способност за самостални научни рад. Изградио је систематичну и критичку анализу постојећих решења, уз уочавање њихових недостатака. Развио је оригиналну методологију која ефикасно решава проблеме реализације СИПП система. Показао је да је методологија практично употребљива њеном применом у оквиру реализованог система. Резултате својих истраживања објавио је у истакнутим часописима од међународног значаја са признатим фактором утицаја.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Допринос изложене докторске дисертације је у домену анализе и синтезе једног оваквог софтверског система, који треба да омогући интелигентно прослеђивање питања написаних на природном језику. Као саставни делови дисертације садржани су следећи научни доприноси:

- Идентификација домена истраживања СИПП система, као и генерални преглед области вештачке интелигенције везане за одговарање на питања написана на природном језику, проналажење експерата и семантичког рутирања упита.
- Систематизација и класификација постојећих решења и генерализација њихових функционалности са становишта унапред дефинисаног СИПП процеса.
- Увођење оригиналне презентационе парадигме која генерализује суштину свих расположивих СИПП решења пронађених у отвореној литератури и која омогућава упоредну анализу и евалуацију оваквих система.
- Формирање методологије пројектовања СИПП система на основу класификације, анализе и евалуације решења која су била примењена, која се примењују или која се могу применити у оквиру ове дисциплине.
- На основу изведених закључака и уочених проблема дат је предлог и имплементација новог софтверског система који треба да омогући интелигентно прослеђивање питања написаних на природном језику.
- У оквиру предложеног приступа реализован је нови приступ за обраду питања, који омогућава њихову визуализацију, што обезбеђује интуитивну представу специфичних односа између концепата, као и њиховог значаја у питању. Такође, овај приступ комбинује потпуно аутоматску обраду текста и ручну корекцију резултата, пружајући кориснику могућност повећања тачности излаза. Истовремено, реализовани модул за обраду текста употребљен је и за анализу одговора.
- Анализирани су и дискутовани постојећи приступи за одређивање семантичке сличности два кратка текста, погодни за језике са врло ограниченим електронским лингвистичким ресурсима, где је посебан акценат стављен на српски језик.
- На основу донетих закључака предложен је нови алгоритам, назван LInSTSS, који приликом одређивања семантичке сличности два кратка текста узима у обзир и специфичности речи које ови текстови садрже. Такође, реализован је корпус парафраза за српски језик над којим је извршена евалуација. Резултати добијени над овим корпусом показали су да предложени алгоритам пружа боље резултате у односу на постојећа решења.

- У оквиру реализације фазе прослеђивања питања, дискутоване су специфичности проблема поређења питања и корисничких профила, и предложен је нови алгоритам назван P2Q. Добијени резултати показали су да овај приступ пружа знатно боље резултате у односу на остале евалуиране приступе.
- Дат је преглед и анализа доступних веб портала чији подаци се могу употребити за формирање корпуса питања и одговора. На основу извршене анализе одабран је један помоћу кога је формиран корпус, а који је затим употребљен за евалуацију целокупног система и тестирање полазних хипотеза.

4.2. Критичка анализа резултата истраживања

Увидом у дисертацију, полазне хипотезе и циљеве истраживања, Комисија констатује да је кандидат успешно одговорио на постављене изазове, и да резултати оправдавају почетна очекивања. Предложен је оригиналан приступ развоју система за интелигентно прослеђивање питања написаних на природном језику. Реализација алгоритма семантичке сличности два кратка текста која узима у обзир и специфичности речи које ови текстови садрже и алгоритма за решавање проблема поређења питања и корисничких профила изузетно је добило на значају последњих година развојем интернета и система за колаборацију. Провера практичне употребљивости ових метода у оквиру реализованог СИПП система потврђује да предложени приступ нема само теоријски значај. Особине приступа чине га јединственим у односу на конкурентске приступе и употребљивим у великом броју реалних примена. Резултати и објашњења истраживања из дисертације су од интереса за све оне који желе да се баве овим новим научим подручјем, да се упознају са појмовима и концептима и да употребе предложену методологију за реализацију софтверског СИПП система. Такође, дисертација може допринети покретању нових веб портала овога типа, као и унапређењу размене знања уопштено. Коначно, дисертација може буде од користи за будуће генерације студената докторских студија, инжењере, практичаре и истраживаче које интересује ова област и који желе да дају свој допринос.

4.3. Верификација научних доприноса

Кандидат је објавио следеће радове који су у непосредној вези са докторском дисертацијом:

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*, ISSN 0167-9236, Vol. 55, Issue 3, pp. 710–719, June 2013. (**IF(2012) 2.201, IF-5(2012) 3.037**), doi: 10.1016/j.dss.2013.02.002

Категорија M22:

1. **Furlan B.,** Nikolić B., Milutinović V., "A Survey and Evaluation of State-of-the-Art Intelligent Question Routing Systems," *International Journal of Intelligent Systems*, ISSN 0884-8173, Vol. 28, Issue 7, pp. 686–708, July 2013., (**IF(2012) 1.416, IF-5(2012) 1.579**) doi: 10.1002/int.21597

Категорија M53:

1. **Furlan B.,** Sivački V., Jovanović D., Nikolić B. "Comparable Evaluation of Contemporary Corpus-Based and Knowledge-Based Semantic Similarity Measures of Short Texts," *JITA*, vol. 1, no. 1, ISSN 2233-0194 (online), pp. 65-71, June 2011. (**nema impakt faktor**)

2. Varga E., **Furlan B.**, and Milutinovic V., "Document Filter Based on Extracted Concepts," *Transactions on Internet Research*, vol. 6, no. 1, ISSN 1820 – 4503 (online), pp. 5-9, January 2010. (**nema impakt faktor**)

Kategorija M33:

1. **Furlan B.**, Žitnik S., Nikolić B., Bajec M., "The Role of Semantic Similarity for Intelligent Question Routing," in Informatics, Spiska, Nova Ves, Slovakia, November 5th – 7th 2013.
2. **Furlan B.**, Nikolic B., Milutinovic V., "A Survey of Intelligent Question Routing Systems," in IEEE Intelligent Systems, Sofia, Bulgaria, September 2012.
3. Jelisavčić V., **Furlan B.**, Protić J., Milutinović V., "Topic Models and Advanced Algorithms for Profiling of Knowledge in Scientific Papers," in MIPRO, Opatija, Croatia, May 2012.
4. Žitnik S., Subelj L., Jankovic M., **Furlan B.**, Draskovic D., Kojic N., Misic M., Bajec M., "Iterative End-to-end Information Extraction based on Linear Models," in ERK, Portorož, Slovenia, September 2013.
5. Stojanović S., **Furlan B.**, Tomašević M., Milutinović V., Fellow of the IEEE, "An Overview of Concurrency Support in Accessing Shared Data in SMPs," In ACACE, L'Aquila, Italy, 2008.

Kategorija M34:

1. **Furlan B.**, Nikolic B., Milutinovic V., "Intelligent Question Routing: An Overview of Some Recent Advances and Open Problems", VIPSI, Miločer, Crna Gora, 2011.
2. **Furlan B.**, "An Intelligent Question Routing System," VIPSI, Pisa, Italy, 2008.

Kategorija M63:

1. **Furlan B.**, Stamenković J., Nikolić B., Mišić M., "Algoritam određivanja semantičke sličnosti između korisničkog profila i pitanja," ETRAN, Zlatibor, Srbija, 3 - 6. Juna 2013.
2. Jelisavčić V., **Furlan B.**, Protić J., Milutinović V., "Knowledge Modeling and Classification of Scientific Papers Based on Topic Modeling," in YUINFO, Kopaonik, Serbia, March 2012. pp. 664-669
3. Batanović V., **Furlan B.**, Nikolić B., "Softverski sistem za određivanje semantičke sličnosti kratkih tekstova na srpskom jeziku," TELFOR, Beograd, Srbija, 22-24. Novembra, 2011.
4. Jovanović D., **Furlan B.**, Nikolić B., "Softverski sistem za automatsko određivanje semantičke sličnosti kratkog teksta," ETRAN, Banja Vrućica (Teslić), R. Srpska, BIH, 6-9. Juna, 2011.
5. **Furlan B.**, Nikolić B., "Veb-Baziran Sistem za Efikasno Dobijanje Odgovora," ETRAN, Palić, Srbija, 2008.
6. Nikolić S., **Furlan B.**, Josipović P., "aLive! - Sistem za inteligentno prosleđivanje pitanja," YUINFO, Kopaonik, Srbija, 2008.

5. ЗАКЉУЧАК И ПРЕДЛОГ

Дисертација кандидата дипл. инж. Бојана Фурлана, под насловом „Методологија пројектовања система за интелигентно прослеђивање питања написаних на природном језику“ представља оригиналан, савремен и значајан научни допринос. Текст дисертације написан је јасно и разумљиво и добро је организован кроз поглавља и одељке. Циљеви дисертације су јасно формулисани и мотивисани, а резултати истраживања систематски изложени, тако да се научни доприноси могу недвосмислено утврдити. У спроведеним истраживањима предложена је оригинална методологија развоја система за интелигентно прослеђивање питања написаних на природном језику. Извршена је реализација два нова алгорита, и то алгорита за одређивање семантичке сличности два кратка текста и алгорита за решавање проблема поређења питања и корисничких профила. Практична примена ових алгорита је потврђена у оквиру реализованог система за интелигентно прослеђивање питања написаних на природном језику. Објављивањем резултата својих истраживања у истакнутим часописима од међународног значаја, кандидат је показао способност за самосталан научни рад, а доприноси истраживања добили адекватну потврду ваљаности.

Комисија констатује да дисертација садржи оригиналне научне доприносе, испуњава све законске, формалне и суштинске услове, као и све критеријуме који се уобичајено примењују приликом вредновања докторских дисертација на Електротехничком факултету у Београду. Комисија са задовољством предлаже Наставно-научном већу Електротехничког факултета у Београду да се докторска дисертација под називом „Методологија пројектовања система за интелигентно прослеђивање питања написаних на природном језику“ кандидата дипл. инж. Бојана Фурлана прихвати, а кандидату одобри усмена одбрана.

У Београду, 08.11.2013.

ЧЛАНОВИ КОМИСИЈЕ

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

др Вељко Милутиновић, редовни професор
Универзитет у Београду – Електротехнички факултет

др Ненад Митић, ванредни професор
Универзитет у Београду – Математички факултет

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

др Милош Цветановић, доцент
Универзитет у Београду – Електротехнички факултет