

28.02.2014. године
Београд,
04-03-11/44

УНИВЕРЗИТЕТ У БЕОГРАДУ
Веће научних области техничких
наука
Београд
Студентски трг бр 1

У прилогу дописа достављамо извештај комисије о оцени завршене докторске дисертације кандидата Милана Вукићевића, са пратећом документацијом.

Докторска дисертација и извештај комисије о оцени завршене докторске дисертације кандидата Милана Вукићевића стављени су на увид јавности и у предвиђеном року није било примедби на изложену докторску дисертацију.

С поштовањем,

Секретар факултета

Гордана Алексић Славински, дипл. правник

Факултет организационих наука

УНИВЕРЗИТЕТ У БЕОГРАДУ

04-03-11/44

(Број индекса)

Веће научних области техничких наука
(назив стручног већа коме се захтев упућује, сходно
чл 75. Статута Универзитета у Београду и
чл.7. ст. 1. овог Правилника)

28.02.2014. године

(Датум)

Београд

Студентски трг

бр.1

ЗАХТЕВ
за давање сагласности на извештај о урађеној докторској дисертацији

Молимо да, сходно члану 68. ст. 3. Закона о универзитету («Службени гласник РС» бр. 20/98), дате сагласности на Извештај о урађеној докторској дисертацији кандидата

Милана (Жарко) Вукићевића

(име, име једног родитеља и презиме)

Кандидат Милан (Жарко) Вукићевић пријавио је докторску дисертацију под називом:

(име, име једног родитеља и презиме)

Развој и пројектовање алгоритама за кластеровање експресије гена

Универзитет је дана 16.09.2013. године својим актом под 61206-3974/2-13 дао сагласност на предлог теме докторске дисертације која је гласила

Развој и пројектовање алгоритама за кластеровање експресије гена

Комисија за оцену и одбрану докторске дисертације кандидата Милана (Жарко) Вукићевића

(име, име једног родитеља и презиме)

образована је на седници одржаној 25.12.2013. године, одлуком Факултета под 3/184-9 у саставу:

име и презиме члана комисије	звање	научна област
1. др Милија Сукновић	ред. проф. Факултета организационих наука	моделирање пословних система и пословно одлучивање
2. др Борис Делибашић	ванр. проф. Факултета организационих наука	моделирање пословних система и пословно одлучивање
3. др Драган Радојевић	научни саветник Института Михајло Пупин	меко рачунарство и рачунарска интелигенција

Наставно-научно веће Факултета прихватило је извештај Комисије за оцену и одбрану докторске дисертације на седници одржаној дана 12.02.2014. године.

ДЕКАН ФАКУЛТЕТА

Прилог:

1. Извештај Комисије са предлогом,
2. Акт Наставно-научног већа Факултета о усвајању Извештаја,
3. Примедбе дате у току стављања Извештаја на увид јавности, уколико је таквих примедби било.

Проф. др Милан Мартић

05-01 бр 3/10-7
18.2.2014.

Поводом предлога Већа студијског програма докторских академских студија, а на основу члана 56. Статута Факултета, на седници Наставно-научног већа одржаној 12.2.2014. године, једногласно је донета следећа

ОДЛУКА

Прихвата се Извештај Комисије за оцену завршене докторске дисертације кандидата Милана Вукићевића, под насловом »РАЗВОЈ И ПРОЈЕКТОВАЊЕ АЛГОРИТАМА ЗА КЛАСТЕРОВАЊЕ ЕКСПРЕСИЈА ГЕНА« и даје се на увид јавности.

На већу научних области техничких наука Универзитета у Београду од 16.09.2013. године број: 61206-3974/2-13 одобрена је израда докторске дисертације кандидата Милана Вукићевића под називом »РАЗВОЈ И ПРОЈЕКТОВАЊЕ АЛГОРИТАМА ЗА КЛАСТЕРОВАЊЕ ЕКСПРЕСИЈА ГЕНА«

Кандидат Милан Вукићевић је објавио следећи рад који га квалификује да може да одбрани докторску дисертацију:

1. Vukićević M., Jovanović M., Delibasic B., Išlamović S., Suknović M. (2012) Reusable component-based architecture for decision tree algorithm design, International Journal on Artificial Intelligence Tools 5 (3), pp. 597-610, ISSN: 1793-6349 doi: <http://dx.doi.org/10.1142/S0218213012500224> IF 0.250

Председник Наставно научног већа

Проф. др Милан Мартић

Доставити:

- Стручном сараднику за ПДС
- Секретару органа управљања и стручних органа
- Архиви

НАУЧНО-НАСТАВНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Милана Вукићевића

Одлуком бр. 3/184-9 од 25.12.2013. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Милана Вукићевића под насловом

Развој и пројектовање алгоритама за кластеровање експресија гена

После прегледа достављене Дисертације и других пратећих материјала и разговора са Кандидатом Миланом Вукићевићем, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Тему приступног рада Милан Вукићевић је пријавио 05.07.2013. Комисија за оцену научне заснованости пријављене докторске дисертације формирана је 10.07.2013. бр 3/109-14. Кандидат је приступни рад одбранио 21.08.2013. Одлука о усвајању извештаја Комисије о научној заснованости пријављене докторске дисертације донета је 28.08.2013. бр 3/119-7. Одлуком Универзитета у Београду од 16.09.2013. бр 61206-3974/2-13 даје се сагласност на предлог теме докторске дисертације Милана Вукићевића под називом „Развој и пројектовање алгоритама за кластеровање експресија гена“, а за ментора је именован проф. др Милија Сукновић, редовни професор Факултета Организационих Наука, Универзитета у Београду. Ментор је известио Наставно научно веће Факултета Организационих Наука да је Милан Вукићевић завршио израду докторске дисертације, а Наставно научно веће Факултета Организационих Наука је именовало Комисију за оцену завршене докторске дисертације 25.12.2013. бр 3/184-9.

1.2. Научна област дисертације

Предмет дисертације је предлог нове методологије за развој, пројектовање и селекцију алгорита за кластеровање података о експресији гена. Централни проблем је пројектовање и/или идентификација алгоритама кластеровања који имају добре перформансе на високодимензионалним подацима о експресији гена.

Докторска дисертација припада области техничких наука и подручју организационих наука. Ужа научна област којом се бави дисертација је Моделирање пословних система и пословно одлучивање. Ментор дисертације је

проф. др Милија Сукновић, редовни професор Факултета Организационих Наука, Универзитета у Београду.

1.4 Биографски подаци о кандидату

Милан Вукићевић је рођен 12.06.1983. године у Београду, Република Србија. Од тада живи у Београду, где је завршио основно образовање. Средње образовање завршава такође у Београду, у XIII београдској гимназији 2002. године.

Факултет Организационих Наука, Универзитета у Београду, уписује 2002. године. Дипломира на смеру за информационе системе 2007. Дипломске (Мастер) студије на одсеку за Информационе системе и технологије, на Факултету организационих наука, Универзитета у Београду завршио је 2008. године са просечном оценом 10, одбраном дипломског рада под називом „Пројектовање складишта података студентске службе ФОН-а“ чији се практични резултати активно користе при анализи наставног процеса Факултета Организационих наука. Исте године запошљава се као Аналитичар података студентске службе ФОН-а и као демонстратор учествује у настави из предмета Системи за подршку одлучивању и Експертни системи, код проф. др Милије Сукновића. 2008. године држао курсеве и припремио материјале (скрипте и тестове) “Складишта података” и “ОЛАП” у оквиру обуке запослених у Републичком заводу за статистику Србије за фирму Белит. Исте године ради на развоју Апликација за електронску огласну таблу за оглашавање јавних набавки“, за потребе Управе за јавне набавке. За свој наставни рад континуирано је оцењиван од стране студената просечном оценом преко 4,40 на скали до 5.

Запослење

2007 - 2008: Пројектант информационих система, БЕЛИТ - Белграде Информацион Технологиес.

2008 - 2010: Сарадник у настави, Универзитет у Београду, Факултет организационих наука, Катедра за Организацију пословних система.

2010 - 2014: Асистент, Универзитет у Београду, Факултет организационих наука, Катедра за Организацију пословних система.

Континуирана едукација

2009: Похађао курс "Академске вештине" које је организовало Министарство Науке републике Србије а спровео др Steve A. Quarrie

2011: Certified RapidMiner "Analyst"

2012: Certified RapidMiner "Expert"

Учесће у пројектима

1. Инфраструктура за технолошки подржано учење у Србији, научно-истраживачки пројекат у оквиру програма технолошког развоја Министарства науке Републике Србије, Београд, 2011-2014. Број пројекта ИИИ 47003
2. Интеракција етиопатогенетских механизма парадонтопатије и периимплантитиса са системским болестима данашњице, научно-истраживачки пројекат у оквиру програма технолошког развоја Министарства науке Републике Србије, Београд, 2011-2014. Број пројекта ИИИ 47008.
3. Развој платформе за откривање законитости у подацима засноване на генеричким компонентама и приступу „белих кутија“, научно-истраживачки

- пројекат у оквиру програма билатералне сарадње између Министарстава за науку Републике Србије и Републике Немачке - ДААД 2011.
4. Развој платформе за откривање законитости у подацима засноване на генеричким компонентама”, научно-истраживачки пројекат у оквиру програма билатералне сарадње између Министарстава за науку Републике Србије и Републике Немачке - ДААД 2010.
 5. Развој платформе за моделовање компоненти и документовање развијених модела унутар стандардног процеса дејта мајнинга, научно-истраживачки пројекат у оквиру Програма технолошког развоја Министарства науке Републике Србије, Београд, 2008 - 2009. Број пројекта је: TP12013.
 6. Апликација за електронску огласну таблу за оглашавање јавних набавки, за потребе Управе за јавне набавке, Београд, 2007-2008.

Наставно искуство

Основне студије: Теорија одлучивања, Системи за подршку одлучивању, Пословна интелигенција, Дејта Мајнинг, Менаџерско одлучивање.

Мастер студије: Системи пословне интелигенције, Складишта података, Машинско учење.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација кандидата Милана Вукићевића, под насловом „Развој и пројектовање алгоритама за кластеровање експресија гена” написана је на 94 стране и структурирана у 6 поглавља, као и дела који се односи на коришћену литературу.

Садржај

1	Увод.....	8
2	Кластеровање (груписање података).....	9
2.1	Кластеровање експресија гена.....	11
2.2	Алгоритми за кластеровање експресија гена.....	14
2.2.1	Кластеринг базиран на представницима.....	14
2.2.2	Хијерархијски кластеринг.....	18
2.3	Остали алгоритми за кластеровање експресија гена.....	21
2.4	Евалуација алгоритама кластеровања.....	23
2.4.1	Интерне мере евалуације.....	24
2.4.2	Екстерне мере евалуације.....	33
3	Компонентни приступ у развоју алгоритама кластеровања.....	36
3.1	Под-проблеми и поново употребљиве компоненте.....	37
3.2	Генерички алгоритам за кластеровање.....	39
3.2.1	Комплексност генеричког алгоритма.....	43
3.3	Софтверска архитектура за колаборативни дизајн алгоритама	44
3.3.1	Генеричка архитектура базирана на компонентама.....	44
3.3.2	Структура генеричког алгоритма кластеровања.....	46
4	Примена, рангирање и селекција компонентних алгоритама за кластеровање експресија гена.....	48

4.1	Примена партитивног алгоритма кластеровања базираног на компонентама.....	49
4.1.1	Експериментална поставка.....	49
4.1.2	Подаци.....	49
4.1.3	Поређење са добро познатим алгоритмима.....	51
4.1.4	Поређење са резултатима из литературе.....	52
4.1.5	Идентификација добрих компонената за кластеровање података о експресији гена.....	54
4.2	Метод за дизајн алгоритама и оптимизацију броја кластера код кластеровања података о експресији гена.....	60
4.2.1	Подаци.....	60
4.2.2	Експериментална поставка.....	61
4.2.3	Примена ОЗП техника за идентификацију алгоритма прилагођеним подацима.....	63
4.3	Идентификација корелације између интерних и екстерних мера евалуације алгоритама кластеровања.....	67
4.3.1	Преглед сличних приступа.....	68
4.3.2	Корелација интерних мера са АМИ индексом.....	69
4.3.3	Селекција модела базирана на интерним мерама.....	71
4.3.4	Квалитет селектора модела базиран на интерним мерама евалуације.....	73
5	Мета - учење за алгоритме кластеровања за експресију гена.....	75
5.1	Системи мета-учења.....	75
5.2	Проширени модел мета-учења за кластеровање експресија гена.....	78
5.2.1	Основни процес евалуације и селекције мета-модела.....	80
5.2.2	Процес за аутоматску селекцију мета-атрибута.....	84
5.2.3	Експериментални резултати.....	86
6	Закључак.....	93
6.1	Остваренидопринос.....	93
6.2	Правци даљег истраживања.....	93
7	Литература.....	94

2.2. Кратак приказ појединачних поглавља

У уводном поглављу дат је осврт на опште друштвено-технолошке околности које су довеле до предлога теме. Описана је мотивација аутора за рад у датој области као и основна идеја рада.

Друго поглавље разматра кластеровање као задатак откривања законитости у подацима (енг. *data mining*, у даљем тексту ОЗП) са посебним освртом на преглед типова алгоритама који су успешно примењени у области кластеровања експресија гена. Такође је у овом поглављу посебна пажња обрађена на евалуацију модела кластеровања с обзиром да је евалуација један од највећих проблема у примени алгоритама кластеровања генерално.

Треће поглавље детаљно описује предности компонентног приступа при пројектовању и примени алгоритама ОЗП. У овом поглављу се описује генеричка архитектура за колаборативни дизајн алгоритама кластеровања, репозиторијум компоненти и генеричке алгоритме за дизајн нових алгоритама кластеровања базираних на представницима.

Четврто поглавље описује поступак евалуације алгоритама за кластеровање података о експресији гена и предложене моделе за идентификацију квалитетних алгоритама кластеровања који су базирани на методама ОЗП. Такође су детаљно испитане статистичке значајности разлика у перформансама између различитих алгоритама и компонената алгоритама. Посебна пажња је посвећена испитивању постојања корелација између интерних и екстерних мера евалуације како би се идентификовале интерне мере евалуације које се могу оптимизовати у циљу добијања "правих" структура кластера.

Коначно у петом поглављу се описује предложени систем мета-учења за селекцију и рангирање алгоритама за кластеровање експресија гена. Овај систем представља унапређење постојећих система мета-учења за кластеровање експресија гена који укључује компоненте алгоритама кластеровања и интерне мере евалуације као мета- атрибуте. У оквиру овог система развијен је и процес за аутоматско одређивање важности мета-атрибута и тиме побољшава укупне перформансе система и омогућава аутоматско прилагођавање система у случају додавања нових скупова података, алгоритама или мета атрибута. Такође је приказана исцрпна евалуација система и идентификација најважнијих мета-атрибута.

У шестом поглављу се сумирају резултати истраживања, даје се критички осврт на предложено решење, као и правци даљег истраживања.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Тема којом се бави докторска дисертација је савремена и веома атрактивна, како у научном тако и у практичном смислу. Истраживање и развој новог поступка за пројектовање, селекцију и рангирање алгоритама за кластеровање експресија гена има врло широк друштвени значај који се огледа у могућности примене развијеног решења како за дијагностичке потребе, тако и за даљи развој алгоритама кластеровања у другим областима.

У докторској дисертацији је предложен потпуно нов поступак за пројектовање, селекцију и рангирање алгоритама за кластеровање експресија гена, који користи оригиналан приступ проблему и као такав је оригиналан и јединствен.

3.2. Осврт на референтну и коришћену литературу

У раду је коришћена обимна и релевантна међународна литература. Највећа пажња посвећена је публикацијама највећих научних удружења у области информационих технологија, IEEE и ACM Велики број коришћених извора публикован је на водећим конференцијама и у водећим часописима из области. Поред тога коришћена је обимна и актуелна литература коју су публиковали водећи светски издавачи (Springer, Elsevier, Wiley, Prentice-Hall, Taylor&Francis).

У изради овог рада референцирано је 166 извора релевантних за обраду ове теме.

3.3. Опис и адекватност примењених научних метода

У изради дисертације коришћене су следеће научне методе:

У првом делу дисертације (поглавља 1 и 2) коришћене су методе прикупљања и анализе постојећих научних резултата и достигнућа као и критичка анализа постојећих приступа.

За развој нове методологије коришћене су модерне као и стандардне методе евалуације кластер модела. За идентификацију значајности разлика између алгоритама и/или њихових компонента коришћени су традиционални статистички тестови. Такође, модели откривања законитости у подацима (класификација, асоцијација и процена) су евалуирани стандардним методама евалуације.

Начин на који је коришћена литература, њени извори и критички осврт јасно указују на адекватан истраживачки приступ. У реализацији софтверских решења коришћена је објектно-оријентисана методологија пројектовања употребом UML језика за моделовање. Експериментална евалуација поступка је адекватно постављена и извршена ослањајући се на водеће стандарде за евалуацију модела кластеровања.

На основу анализе докторске дисертације, може се закључити да примењене научне методе и технике одговарају, по свом значају и структури, теми дисертације и спроведеном истраживању.

3.4. Применљивост остварених резултата

Успешност предложеног поступка проверена је низом емпиријских евалуација. Резултати су показали да хипотеза: *могуће је пројектовати алгоритме кластеровања експресије гена, који ће бити бољи од постојећих решења и могуће је идентификовати најбољи алгоритам за конкретан скуп података*, не може бити одбачена. Алгоритми кластеровања базирани на компонентама су показали боље перформансе у односу на оригиналне алгоритме из којих су генерисани и упоредиве перформансе са другим типовима кластер алгоритама. Предложени систем за селекцију и рангирање алгоритама кластеровања се показао као врло прецизан у процени перформанси алгоритама.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат је у досадашњем научно истраживачком раду показао изузетну способност и интуитивност. Истакао се учешћем у водећим светским и националним конференцијама које су као фокус имале област дисертације. Током рада на дисертацији објавио је као аутор или коаутор 38 број радова у водећим међународним и националним часописима, националним и међународним конференцијама и поглављима књига. Као истраживач, учествовао је у неколико домаћих научно-истраживачких пројеката као и у два међународна пројекта. Такође током периода докторских студија имао је прилике да менторише и усмерава научно истраживачки рад млађих колега са којима је у сарадњи остварио високе резултате.

Узевши у обзир целокупно залагање и остварене резултате у научно истраживачком раду, закључујемо да је кандидат способан да се у потпуности самостално бави научно-истраживачким радом.

4. ОСТВАРЕН НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Континуирано истраживање у области развоја компонентних алгоритама кластеровања и њихове примене на подацима о експресијама гена. Најзначајнији допринос је предлог оригиналне методологије за пројектовање, евалуацију и селекцију алгоритама за кластеровање података о експресији гена која је заснована на проширеном систему мета-учења, који интегрише особине алгоритама (компоненте) и интерне мере евалуације као мета-атрибуте. Овај систем такође укључује процес за аутоматску селекцију мета-модела базиран на аутоматској селекцији мета-атрибута чиме омогућава једноставно ажурирање и прилагођавање новим скуповима података, алгоритмима и мета подацима.

Поред тога, научни допринос рада огледа се и у следећем:

- Преглед савремених модела који могу да се користе за кластеровање података о експресијама гена.
- Предлог архитектуре за колаборативни дизајн компонентних алгоритама кластеровања.
- Предлог генеричких алгоритама кластеровања (дивизиони и хијерархијско-дивизиони).
- Примена и детаљна анализа компонентних алгоритама кластеровања на реалним подацима као и поређење са другим алгоритмима из литературе.
- Предлог метода за идентификацију компонената које учествују у дизајну алгоритама са добрим перформансама и које би требало разматрати при дизајну нових алгоритама.
- Предлог модела мета-учења прилагођеног алгоритмима за кластеровање података о експресији гена.

4.2. Критичка анализа резултата истраживања

Реализовани резултати истраживања представљају целовит преглед области кластеровања података о експресијама гена и новопредложена методологија за развој, евалуацију, селекцију и рангирање алгоритама кластеровања. Примена дефинисаног поступка је проверена реализацијом прототипа модела заснованог на техникама ОЗП као и модела мета-учења. Успешност предложене методологије је проверена емпиријски, применом различитих експеримената на прикупљеној бази података о експресији гена задовољавајуће величине.

4.3. Верификација научних доприноса

Области научног интересовања Милана Вукићевића су откривање законитости у подацима (енг. *data mining*), машинско учење, системи пословне интелигенције, системи за подршку одлучивања, складишта података и теорија одлучивања. Током свог научно истраживачког рада објавио је велики број научних радова и то у следећим категоријама: M13 – три рада, M21 – два рада, M22 – четири рада, M23 – седам радова, M33 – седам радова, M34 - три рада, M51 – један рад, M53 - 1 рад, M63 – десет радова.

У наставку је дат је одабрани подскуп радова у којима су директно или индиректно описани кључни доприноси из теме ове дисертације разврстани по категоријама.

Категорија M21:

1. **M. Vukićević**, S. Radovanović, M. Milovanović, and M. Minović, Cloud Based Meta-learning System for Predictive Modeling of Biomedical Data, Scientific World Journal, *in press* (Accepted 21 January 2014), IF (2012) 1.730.
2. **Vukićević M**, Kirchner K, Delibasic B, Jovanovic M, Ruhland J, Suknović M (2012) Finding best algorithmic components for clustering microarray data, Knowledge and Information Systems, <http://dx.doi.org/10.1007/s10115-012-0542-5>, ISSN: 0219-1377, IF (2011) 2.225.

Категорија M22:

1. Delibašić B., **Vukićević M.**, Jovanović M., Kirchner K., Ruhland J., Suknović M. (2012) An architecture for component-based design of representative-based clustering algorithms, Data & Knowledge Engineering. doi: 10.1016/j.datak.2012.03.005 IF (2012) 1.519.
2. Jovanović M., **Vukićević M.**, Milovanović M., Minović M. (2012), Using data mining on student behavior and cognitive style data for improving e-learning systems: a case study, International Journal of Computational Intelligence Systems 5(3), 597-610, ISSN: 1875-6891, IF (2010) 1.471.
3. Delibašić B., **Vukićević M.** and Jovanović M. (2013) White-Box Decision Tree Algorithms: A Pilot Study on Perceived Usefulness, Perceived Ease of Use, and Perceived Understanding, International Journal of Engineering Education 29 (3), pp. 674-687, IF (2012) 1.925.
4. Delibašić B., **Vukićević M.**, Jovanović M., Suknović M. (2012) White-Box or Black-Box Decision Tree Algorithms: Which to Use in Education?, IEEE Transactions on Education DOI: <http://dx.doi.org/10.1109/TE.2012.2217342>, ISSN: 0169-023X, IF (2012) 0.950.

Категорија M23:

1. Delibašić B., Kirchner K., Ruhland J., Jovanović M., **Vukićević M.** (2009) Reusable components for partitioning clustering algorithms, Artificial Intelligence Review 32 (1), 59-75. <http://dx.doi.org/10.1007/s10462-009-9133-6>, ISSN: 0269-2821, IF (2009) 0.057.
2. Suknović M., Delibašić B., Jovanović M., **Vukićević M.**, Becejski-Vujaklija D., Obradović Z. (2011) Reusable Components in Decision Tree Induction Algorithms, Computational Statistics. DOI: <http://dx.doi.org/10.1007/s00180-011-0242-8>, ISSN: 1613-9658, IF(2011) 0.276.
3. Delibašić B., Jovanović M., **Vukićević M.**, Suknović M., Obradović Z. (2011) Component-based decision trees for classification, Intelligent Data Analysis 15(5), 671-693, <http://dx.doi.org/10.3233/IDA-2011-0489>, ISSN: 1088-467X, IF (2011) 0.448.
4. **Vukićević M.**, Jovanović M., Delibašić B., Išljamović S., Suknović M. (2012) Reusable component-based architecture for decision tree algorithm design, International Journal on Artificial Intelligence Tools. doi:10.1142/S0218213012500224, ISSN: 0218-2130, IF (2012) 0.250.

Категорија M13:

1. **Vukićević, M.**, Jovanović, M., Delibašić, B., & Suknović, M. (2013). Grouping Higher Education Students with RapidMiner. RapidMiner: Data Mining Use Cases and Business Analytics Applications, Eds. Markus Hofmann, Ralf Klinkenberg, Chapman & Hall/CRC Data Mining and Knowledge Discovery Series.
2. **Vukićević, M.**, Jovanović, M., Delibašić, B., & Suknović, M. (2013). Recommender System for Selection of the Right Study Program for Higher Education Students. RapidMiner: Data Mining Use Cases and Business Analytics Applications, Eds. Markus Hofmann, Ralf Klinkenberg, Chapman & Hall/CRC Data Mining and Knowledge Discovery Series.
3. Jovanović M., **Vukićević M.**, Delibašić M., Suknović M. (2013). Using RapidMiner for Research: Experimental Evaluation of Learners. RapidMiner: Data Mining Use Cases and Business Analytics Applications, Eds. Markus Hofmann, Ralf Klinkenberg, Chapman & Hall/CRC Data Mining and Knowledge Discovery Series.

Категорија M33:

1. **Vukićević M.**, Jovanović M., Delibašić B., Suknović M. (2010) WhiBo - RapidMiner plug-in for component based data mining algorithm design, In proc. of the 1st RapidMiner Community Meeting and Conference, September 13-16, Dortmund, Germany, www.rcomm2010.org, p. 30-35.
2. Jovanović M., Delibašić B., **Vukićević M.**, Suknović M. (2011) Optimizing performance of decision tree component-based algorithms using evolutionary algorithm in RapidMiner, In proc. of the 2nd RapidMiner Community Meeting and Conference, June 7-10, Dublin, Ireland, www.rcomm2011.org, 135-149, ISBN 978-3-8440-0093-1
3. **Vukićević M.**, Delibašić B., Jovanović M, Suknović M, Obradović Z (2011) Internal Evaluation Measures as Proxies for External Indices in Clustering Gene Expression Data, In proc. of the 2011 IEEE International Conference on Bioinformatics and Biomedicine (BIBM11), Atlanta, Georgia, USA, Nov. 12-15, 574-577
4. **Vukićević M.**, Delibašić B., Obradović, Z., Jovanović M., Suknović, M. (2012) "A Method for Design of Data-tailored Partitioning Algorithms for Optimizing the Number of Clusters in Microarray Analysis," Proc. 2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, San Diego, CA, May 2012.
5. Jovanović M., Stojanović J., **Vukićević M.**, Stojanovic V., Delibašić B., Suknović M., (2012) NeurophRM: Integration of the Neuroph framework into RapidMiner, In proc. of the 3rd RapidMiner Community Meeting and Conference, August 28- September 1, Budapest, Hungary, www.rcomm2012.org
6. Radovanović S., **Vukićević M.**, Jovanović M., Delibašić B., Suknović M., Meta-learning system for clustering gene expression microarray data In proc. of the 4th RapidMiner Community Meeting and Conference, August 27- August 29, Porto, Portugal, www.rcomm2013.org

Категорија M63:

1. Delibašić B, Suknović M, Bečejski-Vujaklija D, Jovanović M, **Vukićević M** (2008) Patern platforma za dejta majning, SIMOPIS, 287 - 290, ISBN: 978-86-7395-248-2

2. Jovanović M., Suknović M., **Vukićević M.**, Delibašić B. (2009) A white box approach in modeling phase of the data mining process, Proceedings 36th Serbian OR Symposium SIMOPIS, Ivanjica, Sept. 22-25, 709-712, ISBN: 978-86-80953-43-4
3. Jovanović M, Delibašić B., **Vukićević M.**, Suknović M. (2010) An open-source platform for design and testing of data mining algorithms (in Serbian) In Proc. of the 37th serbian operations research conference SYMOPIS, p. 769-772, September 21 - 24, Tara, Serbia, ISBN: 978-86-335-0299-3
4. **Vukićević M.**, Suknović M, Jovanović M (2010) An open source environment for generic partitioning clustering algorithm design, SYMORG 2010, Fakultet organizacionih nauka, Zlatibor, Zbornik radova izdat u CD formatu, ISBN: 978-86-7680-216-6

5. ЗАКЉУЧАК И ПРЕДЛОГ

5.1. Кратак осврт на дисертацију у целини, на научне доприносе конкретној научној области и методологији, на примену у пракси и на способности кандидата

Докторска дисертација се бави значајном и још увек недовољно истраженом области кластеровања експресије гена у циљу дијагнозе или ране идентификације болести или стадијума болести пацијената. Додатни подстицај анализи генских експресија дао је брз развој технологија прикупљања података генских експресија. Технологије ДНК микрочипова омогућавају симултано праћење великог броја карактеристика генских експресија са малом количином грешака. Такође, заменом стандардних дијагностичких алата долази се, првенствено због замене и/или унапређења експертског знања, до смањења трошкова, смањења времена обраде, као и могућности обраде без доменског знања. Због тога ова тема потенцијално има велики утицај на унапређење дијагностичких метода као и на проналажење адекватних терапија за тешке болести и то је чини јако актуелном у областима биоинформатике и откривања законитости у подацима (енг. *data mining*).

Основна идеја овог рада је била да се предложи методологија која ће потенцијално дати решење за овај проблем тако што ће понудити оквир за брзи и колаборативни развој алгорита кластеровања, као и различите моделе за селекцију алгоритама који су најбоље прилагођени подацима о експресији гена. Полазна хипотеза аутора је била: *могуће је пројектовати алгоритме кластеровања експресије гена, који ће бити бољи од постојећих решења и да је могуће је идентификовати најбољи алгоритам за конкретан скуп података*. За ту потребу је развијена архитектура за једноставни дизајн алгоритама кластеровања базираних на компонентама, као и неколико модела за идентификацију најбољих алгоритама за кластеровање експресија гена.

Упоредним прегледом постојећих приступа у обрађеној области дата је тренутна слика академских активности у овој области. На основу анализе објављених истраживања идентификован је простор за предлог нове методологије за дизајн, евалуацију и селекцију алгоритама за кластеровање експресија гена. Ова методологија укључује генерисање великог броја компонентних алгоритама кластеровања, њихову евалуацију над подацима о експресији гена и екстракцију мета-података. Сви подаци (мета-подаци о алгоритмима, мета-поадаци о

алгоритмима, перформансе) се чувају у бази мета-случајева. Над том базом се генерише регресиони модел за селекцију и рангирање алгоритама кластеровања за нове скупове података.

За потребе новог поступка развијена је експериментална платформа која у потпуности подржава предложени процес.

Ради провере исправности постављеног поступка дефинисане су одговарајуће експерименталне провере. Прикупљени су репрезентативни подаци о експресијама гена. Низом експеримената испитане су перформансе развијеног система. На основу добијених резултата закључено је да је предложени поступак научно заснован, да су постигнути одлични резултати у смислу перформанси компонентних алгоритама, као и система за селекцију и рангирање алгорита и да је могуће користити систем у реалном времену.

С обзиром на постигнуте резултате, комплексност, актуелност и специфичност обрађене теме, ова дисертација задовољава највише критеријуме и показује способност колеге Милана Вукићевића за научно-истраживачки рад.

5.2. Предлог Комисије Научно-наставном већу

Предлажемо да се приложена докторска дисертација под називом „Развој и пројектовање алгоритама за кластеровање експресија гена“ кандидата Милана Вукићевића прихвати, изложи на увид у јавности и упуту на коначно усвајање већу техничких наука Универзитета у Београду.

ЧЛАНОВИ КОМИСИЈЕ

Проф. др. Милија Сукновић, редовни професор
Универзитет у Београду, Факултет Организационих Наука

Проф. др. Борис Делибашић, ванредни професор
Универзитет у Београду, Факултет Организационих Наука

Др. Драган Радојевић, научни саветник, Институт Михајло Пупин