

ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСKE ДИСЕРТАЦИЈЕ  
КАНДИДАТА ЗОЛТАНА ГЕЛЕРА

|   |
|---|
| <b>I ПОДАЦИ О КОМИСИЈИ</b>  |
| <p>1. Датум и орган који је именовao комисију<br/>27. фебруар 2015. године, Наставно-научно веће Природно-математичког факултета Универзитета у Новом Саду</p> <p>2. Састав комисије са назнаком имена и презимена сваког члана, звања, назива уже научне области за коју је изабран у звање, датума избора у звање и назив факултета, установе у којој је члан комисије запослен:</p> <ul style="list-style-type: none"><li>- др Милош Радовановић, доцент, <i>Рачунарске науке</i>, 12.12.2011. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – председник</li><li>- др Мирјана Ивановић, редовни професор, <i>Рачунарске науке и информатика</i>, 29. 04. 2002. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – ментор</li><li>- др Зоран Будимац, редовни професор, <i>Рачунарске науке и информатика</i>, 15. 06. 2004. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – члан</li><li>- др Владимир Курбалија, ванредни професор, <i>Рачунарске науке</i>, 01.02.2015. године, Природно-математички факултет, Департман за математику и информатику, Универзитет у Новом Саду – члан</li><li>- др Зоран Боснић, ванредни професор, <i>Рачунарство и информатика</i>, 11.03.2014. године, Факултет за рачунарство и информатику, Универзитет у Љубљани - члан</li></ul> |
| <b>II ПОДАЦИ О КАНДИДАТУ</b>  |
| <p>1. Име, име једног родитеља, презиме:<br/>Золтан (Јован) Гелер</p> <p>2. Датум рођења, општина, држава:<br/>21.10.1978. године, Сомбор, Сомбор, Србија</p> <p>3. Назив факултета, назив студијског програма дипломских академских студија – мастер и стечени стручни назив<br/>Природно-математички факултет, Дипломирани информатичар - мастер, Дипломирани информатичар - мастер</p> <p>4. Година уписа на докторске студије и назив студијског програма докторских студија<br/>2008, Информатика - докторске студије</p> <p>5. Назив факултета, назив магистарске тезе, научна област и датум одбране:</p> <p>6. Научна област из које је стечено академско звање магистра наука:</p>   |
| <b>III НАСЛОВ ДОКТОРСKE ДИСЕРТАЦИЈЕ:</b><br>Role of Similarity Measures in Time Series Analysis<br>(Улога мера сличности у анализи временских серија)   |
| <b>IV ПРЕГЛЕД ДОКТОРСKE ДИСЕРТАЦИЈЕ:</b><br>Навести кратак садржај са назнаком броја страна, поглавља, слика, шема, графикона и сл.<br>Докторска дисертација има 171 страна А4 формата, организована је у 7 поглавља, садржи 51 табелу, 58 слике, 5 прилога и 139 библиографских референци. Поред тога дисертација садржи сажетак на српском језику, кратку биографију кандидата.   |

Дисертација је структурирана на следећи начин:

Chapter 1: Introduction

(Поглавље 1: Увод)

Chapter 2: Time Series and Similarity Measures

(Поглавље 2: Временске серије и мере сличности)

Chapter 3: Time-Series Classification

(Поглавље 3: Класификација временских серија)

Chapter 4: Methods, Tools and Datasets

(Поглавље 4: Методе, алати и скупови података)

Chapter 5: Techniques for Improving Classification Accuracy

(Поглавље 5: Технике за побољшање тачности класификације)

Chapter 6: Framework for Analysis and Prediction (FAP)

(Поглавље 6: Оквир за анализу и предвиђање (FAP))

Chapter 7: Conclusion

(Поглавље 7: Закључак)

Предмет истраживања дисертације обухвата преглед и анализу утицаја глобалних ограничења на најчешће коришћене еластичне мере сличности у области *data mining*-а временских серија са нагласком на тачности класификације. Избор мере сличности један је од најважнијих аспеката анализе временских серија - она треба верно да рефлектује сличност између података приказаних у облику временских серија. Мера сличности представља критичну компоненту многих задатака *mining*-а временских серија, укључујући класификацију, груписање (енг. *clustering*), предвиђање, откривање аномалија и других.

У првом поглављу дат је сажет преглед владајућих ставова и схватања у области истраживања дисертације заједно са мотивацијама и циљевима проучавања приказаних у дисертацији. Неопходни основни појмови *data mining*-а временских серија описани су у другом поглављу: дата је формална дефиниција временских серија, мера сличности и глобалних ограничења проучаваних у оквиру дисертације. Треће поглавље посвећено је класификацији временских серија - детаљно су описане класичне технике класификације: метода најближег суседа (скраћено 1NN) и метода  $k$  најближих суседа (скраћено  $k$ NN), заједно са значајним бројем различитих начина рачунања тежина. Методе, алати и скупови података коришћени у реализацији експеримената описани су у четвртном поглављу.

Детаљан опис опсежних експеримената обухваћених у оквиру истраживања дисертације и дискусија добијених резултата приказани су у петом поглављу рада. Сва испитивања у оквиру ове дисертације изведена су ослањајући се искључиво на FAP (*Framework for Analysis and Prediction*) библиотеку развијену на Департману за математику и информатику Природно-математичког факултета у Новом Саду. Детаљи њене структуре и имплементације представљени су у шестом поглављу.

Последње поглавље закључује дисертацију сумирајући остварене резултате и доприносе и наводи могуће правце за будућа истраживања.

Додаци дисертације садрже графиконе и табеле са детаљним резултатима експеримената који су описани и тумачени у петом поглављу рада.

## **V ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:**

Дисертација у целини, као и њени поједини делови имају добро систематизовану структуру и план излагања. Кандидат је добро систематизовао постојеће радове у области *data mining*-а временских серија са нагласком на улогу мера сличности у класификацији док је својим резултатима дао оригиналан допринос области рачунарских наука. Тиме је у потпуности реализовао постављене циљеве дисертације.

**Наслов.** Наслов дисертације је јасно, прецизно формулисан и одражава садржај

истраживања.

**Поглавље 1: Увод.** У уводном поглављу на јасан и концизан начин су описани основни појмови из области анализе и *mining*-а временских серија заједно са циљевима истраживања. Наведене неформалне дефиниције, праћене релевантном литературом, обезбеђују неопходну основу за разумевање области истраживања дисертације. Јасна мотивација за предмет истраживања, уз указивање на актуелност истог, је растуће интересовање за испитивање различитих задатака анализе временских серија (а пре свега класификације), односно критична улога мера сличности у реализацији тих задатака и широк спектар примене временских серија.

**Поглавље 2: Временске серије и мере сличности.** У другом поглављу дата је формална дефиниција временских серија, мера сличности и глобалних ограничења (*Sakoe-Chiba* појас и *Itakura* паралелограма). Детаљно су описане најчешће коришћене мере сличности које се касније проучавају у оквиру рада: Еуклидско растојање, динамичко искривљење времена (енг. *Dynamic Time Warping* - DTW), најдужи заједнички подниз (енг. *Longest Common Subsequence* - LCS), растојање уређивања са реалном казном (енг. *Edit Distance with Real Penalty* - ERP) и растојање уређивања над реалним серијама (енг. *Edit Distance on Real sequence* - EDR). Сви ови појмови су неопходни за разумевање осталог садржаја дисертације.

**Поглавље 3: Класификација временских серија.** Ово поглавље дисертације посвећено је класификацији временских серија. Објашњена је разлика између класификације (надгледаног учења - енг. *supervised learning*) и груписања (ненадгледаног учења - енг. *unsupervised learning*) и описани су главни кораци процеса класификације. Након прегледа литературе која се бави овом проблематиком, дате су формалне дефиниције класичних техника класификације: методе најближег суседа, методе  $k$  најближих суседа и методе  $k$  најближих суседа са тежинском функцијом. У наставку поглавља дат је опсежан преглед радова који дефинишу различите начине рачунања тежина. У оквиру истраживања дисертације обухваћене су све ове тежинске шеме. Поглавље се завршава описом неколико техника евалуације тачности класификатора коришћених у експериментима описаних у петом поглављу.

**Поглавље 4: Методе, алати и скупови података.** Четврто поглавље описује методе, алате и скупове података који се користе у експериментима чији су резултати описани у петом поглављу. Приказане су основне карактеристике FAP библиотеке и описане су опште особине 46 скупова података UCR (*University of California, Riverside*) репозиторијума временских серија. Ова колекција садржи већину свих јавно доступних скупова означених временских серија у свету. Она се најчешће користи за валидацију различитих концепата *mining*-а временских серија. Обухваћени скупови података потичу из обиља различитих домена, укључујући медицину, роботiku, астрономију, биологију, препознавање лица и рукописа, итд. На крају поглавља укратко су описани наменски алати (као што је генератор матрица растојања) који су развијени за убрзање дуготрајних и рачунски захтевних експеримената обухваћених дисертацијом.

**Поглавље 5: Технике за побољшање тачности класификације.** У петом поглављу приказани су главни резултати дисертације. На основу низа опсежних експеримената истражене су могућности побољшања тачности класификације 1NN и  $k$ NN класификатора применом *Sakoe-Chiba* ограничења на мере сличности и додељивањем различитих тежина најближим суседима. Осим тога, користећи ове технике, извршена је и провера става да је веома тешко надмашити тачност 1 NN класификатора. Након приказа коришћених алгоритама, у првом кораку испитивања истражено је у којој мери се убрзава рачунање сличности између временских серија применама *Sakoe-Chiba* појаса различитих ширина.

У другом кораку експеримената детаљно је испитан утицај примене *Sakoe-Chiba* ограничења на 1 NN класификатор у случају четири најчешће коришћене еластичне мере сличности: DTW, LCS, ERP и EDR. У првој фази експеримената анализиране су промене у графу суседства у односу на смањивање ширине појаса ограничења. У другој фази експеримената истражено је како ове промене утичу на 1 NN класификатор у погледу на ознаке (класе) најближих суседа. Испитивања су заокружена у трећој фази разматрањем утицаја глобалних ограничења на тачност класификације. На основу добијених резултата дате су препоруке за избор оптималне ширине појаса ограничења и

мере сличности које у општем случају дају најбољу тачност класификације.

Испитивања су у трећем кораку проширена и на  $k$ NN класификатор - без и са применом тежинске функције. Да би добили дубљи увид у утицај појаса ограничења, експерименти су укључили неколико различитих метода евалуације тачности класификације. На основу резултата испитивања изводе се закључци о значају првог (најближег) суседа у области класификације временских серија и о могућем утицају примене функције тежине на тачност  $k$ NN класификатора односно на ширину појаса ограничења потребног за постизање најбољих резултата.

У последњем кораку експеримената испитана је могућност побољшања тачности класификације  $k$ NN класификатора применом значајног броја различитих начина рачунања тежина и дата је провера става да је у домену временских серија тешко надмашити методу најближег суседа. На основу резултата експеримената дате су препоруке за избор најуспешнијих техника рачунања тежина.

Циљеви истраживања су јасно формулисани и мотивисани, извођење експеримената је адекватно описано, резултати су приказани прецизно и систематично, закључци су подржани и статистичким тестовима: упареним Вилкоксоним тестом рангова са знаком (енг. *pairwise Wilcoxon sign-rank test*) и са "*corrected resampled*"  $t$ -тестом.

**Поглавље 6: Оквир за анализу и предвиђање (FAP).** Шесто поглавље описује општу структуру и детаље имплементације бесплатне библиотеке отвореног кода FAP, специјализоване за потребе анализе и *mining*-а временских серија. Она је коришћена за реализацију свих експеримената описаних у дисертацији. Дат је прегледан приказ свих главних елемената библиотеке: основне компоненте (тачке, временске серије, скупови временских серија), помоћне класе за праћење и настављање прекинутих дуготрајних процеса израчунавања, мере сличности, класификатори, методе за евалуацију класификатора, технике препроцесирања и различите репрезентације временских серија. На крају поглавља дат је и један пример коришћења FAP библиотеке.

**Поглавље 7: Закључак.** Седмо поглавље даје сажетак резултата истраживања, сумира доприносе дисертације и предлаже могуће правце даљих проучавања.

## VI СПИСАК НАУЧНИХ И СТРУЧНИХ РАДОВА КОЈИ СУ ОБЈАВЉЕНИ ИЛИ ПРИХВАЋЕНИ ЗА ОБЈАВЉИВАЊЕ НА ОСНОВУ РЕЗУЛТАТА ИСТРАЖИВАЊА У ОКВИРУ РАДА НА ДОКТОРСКОЈ ДИСЕРТАЦИЈИ

1. V. Kurbalija, M. Radovanović, **Z. Geler** and M. Ivanović. (2014). The influence of global constraints on similarity measures for time-series databases. *Knowledge-Based Systems* 56:49–67, ISSN: 0950-7051, DOI: <http://dx.doi.org/10.1016/j.knosys.2013.10.021> (M21)
2. D. Mitrović, M. Ivanović and **Z. Geler**. (2014). Agent-based distributed computing for dynamic networks. *Information Technology And Control* 43 (1), 88-97, ISSN: 1392-124X (print), 2335-884X (online), DOI: <http://dx.doi.org/10.5755/j01.itc.43.1.4588> (M23)
3. **Z. Geler**, V. Kurbalija, M. Radovanović, M. Ivanović (2014). Impact of the Sakoe-Chiba band on the DTW time-series distance measure for  $k$ NN classification. In *Knowledge Science, Engineering and Management, 7th International Conference, KSEM 2014, Sibiu, Romania, October 16-18, 2014. Proceedings, Volume 8793*, pp. 105–114, Springer International Publishing, Series ISSN: 0302-9743, ISBN: 978-3-319-12095-9 (print), 978-3-319-12096-6 (online), DOI: 10.1007/978-3-319-12096-6\_10 (M33)
4. D. Mitrović, **Z. Geler**, and M. Ivanović. (2012). Distributed distance matrix generator based on agents. In *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (WIMS 2012)*, ACM, New York, NY, USA, Article 40, 6 pages, ISBN: 978-1-4503-0915-8 DOI: <http://dx.doi.org/10.1145/2254129.2254179> (M33)
5. V. Kurbalija, M. Radovanović, **Z. Geler** and M. Ivanović. (2011). The influence of global constraints on DTW and LCS similarity measures for time-series databases. In *Proceedings of S3T'11, 3rd International Conference on Software, Services and Semantic Technologies, Volume 101 of Advances in Intelligent and Soft Computing*, pages 67–74, Bourgas, Bulgaria,

Springer Berlin Heidelberg (extended version accessible at <http://arxiv.org/abs/1107.0134>) [Best Paper Award] Series ISSN: 1867-5662, ISBN: 978-3-642-23162-9 (print), 978-3-642-23163-6 (online), DOI: 10.1007/978-3-642-23163-6\_10 (M33)

6. V. Kurbalija, M. Radovanović, **Z. Geler** and M. Ivanović. (2010). A framework for time-series analysis. In Proceedings of AIMS'10, 14th International Conference on Artificial Intelligence: Methodology, Systems, Applications, Volume 6304 of Lecture Notes in Artificial Intelligence, pages 42–51, Varna, Bulgaria, Springer Berlin Heidelberg (the framework is accessible at <http://perun.pmf.uns.ac.rs/fap/>) Series ISSN: 0302-9743, ISBN: 978-3-642-15430-0 (print), 978-3-642-15431-7 (online), DOI: 10.1007/978-3-642-15431-7\_5 (M33)

## VII ЗАКЉУЧЦИ ОДНОСНО РЕЗУЛТАТИ ИСТРАЖИВАЊА

На основу резултата низа опсежних експеримената приказаних и тумачених у овој дисертацији, испитане су могућности побољшања тачности класификације 1NN и  $k$ NN класификатора у домену временских серија, ослањајући се на ограничавање прозора искривљења еластичних мера сличности користећи *Sakoe-Chiba* појас, односно на додељивање различитих тежина најближим суседима. Поред тога, користећи ове технике, проверен је и став да је тачност класификације једноставног 1NN класификатора тешко надмашити. Истраживање обухваћено овом дисертацијом усмерено је у неколико праваца:

1. преглед ефеката глобалних ограничења на перформансе рачунања мера сличности,
2. детаљна анализа последице ограничења еластичних мера сличности на тачност класификације класичних техника класификације,
3. опсежна студија утицаја различитих начина рачунања тежина на класификацију временских серија,
4. развој библиотеке отвореног кода (*Framework for Analysis and Prediction - FAP*) која ће да интегрише главне технике и методе потребне за анализу и мининг временских серија и која је коришћена за реализацију експеримената.

Доприноси и резултати остварени и приказани у овој дисертацији су вишеструки:

1. Објашњен је утицај *Sakoe-Chiba* појаса на перформансе ограничених еластичних мера сличности: измерено је време потребно за генерисање матрица растојања за већи број скупова података - за различите вредности параметра  $r$  ограничења. Матрица растојања неког скупа података је матрица у којој елемент на позицији  $(i, j)$  представља растојање између  $i$ -те и  $j$ -те временске серије тог скупа. Израчунавање матрице растојања је дуготрајан процес што га чини погодним за мерење ефикасности глобалних ограничења. За мале величине ограничења ( $r \leq 5\%$ ), разлика у дужини трајања рачунања између неограничених и ограничених верзија мера сличности је величине реда два а негде и три: код неких скупова рачунање је од 10 до скоро 800 пута брже у односу на неограничене мере.
2. Анализирањем графа суседства у односу на промену величине ограничења (параметар  $r$ ) показано је да за мале вредности ограничења (мање од 15%–10% дужине временских серија) ограничене мере постају значајно другачије од неограничених. Поред тога, показано је да ове промене нису исте код различитих мера сличности - DTW је најосетљивија на примену глобалних ограничења, а EDR је најмање осетљива. Пошто резултати класификације у случају 1NN класификатора у потпуности зависе од класе најближег суседа, промене у графу суседства директно утичу на процес класификације. У другој фази експеримената истраживано је у којој мери се мењају најближи суседи класе под утицајем *Sakoe-Chiba* ограничења. Добијени резултати потврдили су сазнања прве фазе испитивања: највеће промене су забележене за DTW, најмање за EDR а LCS и ERP се налазе између ове две мере.
3. Кроз низ исцрпних експеримената показано је да се, у просеку, најбоља тачност класификације постиже за мале вредности параметра  $r$ . Ова вредност је најмања за DTW (око 4% од дужине временских серија) а највећа за ERP (скоро 10%). Промене у графу суседства генерисане са овим вредностима параметра  $r$  су највеће у случају DTW-а (промењено је у просеку око 10%

чворова) а најмањи у случају EDR-а (у просеку, око 1% чворова). Упоредивање тачности класификације 1NN класификатора показано је да DTW генерално има благу предност у односу на остале мере сличности, али експериментални докази нису посебно јаки. За сваку меру сличности можемо наћи барем неколико скупова података за које је дата мера супериорнија у односу на остале. Због тога, избор најбоље мере растојања може варирати од проблема до проблема, без обзира на овај општи резултат.

4. Упоредјујући просечне грешке класификације разматраних еластичних мера сличности (DTW, LCS, ERP и EDR) истакнута је њихова заједничка особина: просечна грешка класификације расте за мале вредности параметра  $r$  ( $<6\%$ ) и достиже свој максимум за  $r=0\%$ . Иако се у општем случају DTW може сматрати најбољим избором, LCS и EDR могу представљати сигурније изборе због мање изражене потребе за пажљиво подешавање параметра  $r$ .
5. Детаљна анализа  $k$  класификатора (за DTW и LCS) је показала да се најбољи резултати, без коришћења тежина, добијају за  $k=1$ . С друге стране, у случају тежинске верзије  $k$ NN класификатора, најбољи резултати се јављају за вредности близу  $k=4$ . Резултати експеримената јасно су потврдили посебан значај првог суседа када је реч о временским серијама. Генерално, можемо закључити да тежинска шема (која даје предност најближем суседу) значајно побољшава тачност класификације за све посматране вредности параметра  $k$  - сва запажања упућују на то да примена тежинске шеме са фаворизовањем најближег суседа може побољшати квалитет и стабилност  $k$ NN класификатора.
6. Посматрајући просечну тачност класификације откривено је да метода  $k$  најближих суседа ( $k$ NN класификатор) - и верзија без тежина, и верзије са тежинама - даје боље резултате од 1NN класификатора у случају значајног броја скупова. У случају свих разматраних мера сличности, најбољи резултати постигнути су помоћу *Dual Dudani* функције тежина. Најлошији резултати добијени су са методом најближег суседа (1NN класификатор) - осим у случају Еуклидског растојања. Важно је напоменути да разлике између најбољих и најлошијих резултата нису нарочито велике ( $<0.01$ ). Резултати статистичких тестова подржавају *Dual Dudani* и *Dudani* шеме рачунања тежина као најбоље изборе.
7. За подршку ових и будућих истраживања развијена је бесплатна библиотека отвореног кода (FAP) која имплементира многе од најважнијих алгоритама у области *mining*-а и анализе временских серија.

Резултати истраживања у оквиру ове дисертације јасно су показали да све главне еластичне мере сличности (DTW, LCS, ERP и EDR) значајно мењају своје понашање за мале вредности глобалног ограничења. Добијени резултати могу помоћи истраживачима у одабиру и подешавању одговарајуће мере сличности у складу са њиховим задацима, што ће убрзати и олакшати и избор и процес подешавања, и обезбедити тачније резултате класификације. Поред тога, увид у понашање мера сличности у односу на мењање ограничења може бити од користи и за креирање ефикасних стратегија индексирања ради налажења (приближно) најближих суседа.

У литератури предложен је већи број различитих начина рачунања тежина најближих суседа. Сваки од ових радова који описује нови начин рачунања тежина извештава о супериорности те нове методе у односу на нека претходна решења, закључци су обично засновани на упоређивању тачности класификације ослањајући се искључиво на Еуклидско растојање, и на основу релативно малог броја скупова података који нису из домена временских серија. У оквиру ове дисертације дата је детаљна анализа утицаја свих ових техника рачунања тежина на тачност класификације  $k$ NN класификатора у области временских серија. Испитивања су обухватила три најчешће коришћене мере сличности временских серија (Еуклидско растојање и неограничене облике DTW-а и LCS-а), а резултати су подржани и статистичким тестовима. Приказани детаљни експерименти потврдили су мишљење да је једноставну методу најближих суседа није лако значајно надмашити у раду са временским серијама, и истакнули *Dual Dudani* и *Dudani* шеме као најбоља решења.

FAP библиотека која је развијена за потребе експеримената у оквиру ове докторске дисертације замишљена је као бесплатан, проширив софтверски пакет отвореног кода који интегрише главне технике и методе неопходне за анализу и *mining* временских серија. Она је пре свега намењена истраживачима временских серија за тестирање и упоређивање постојећих и нових

метода, али ће бити корисна и за практичаре из различитих домена.

#### **VIII ОЦЕНА НАЧИНА ПРИКАЗА И ТУМАЧЕЊА РЕЗУЛТАТА ИСТРАЖИВАЊА**

Експлицитно навести позитивну или негативну оцену начина приказа и тумачења резултата истраживања.

Дисертација је добро структурирана, јасно су формулисани циљеви истраживања, добијени резултати су приказани прегледно и систематично, праћени садржајним табелама и графиконима. Тумачење резултата је коректно изведено уз детаљну анализу, закључци су јасно аргументовани и подржани статистичким тестовима.

#### **IX КОНАЧНА ОЦЕНА ДОКТОРСKE ДИСЕРТАЦИЈЕ:**

Експлицитно навести да ли дисертација јесте или није написана у складу са наведеним образложењем, као и да ли она садржи или не садржи све битне елементе. Дати јасне, прецизне и концизне одговоре на 3. и 4. питање:

1. Да ли је дисертација написана у складу са образложењем наведеним у пријави теме

Дисертација је написана у складу са планом и програмом које је кандидат предложио приликом пријаве теме. Добијени резултати су у складу са постављеним циљевима истраживања.

2. Да ли дисертација садржи све битне елементе

Дисертација садржи све битне елементе: јасно су дефинисани циљеви истраживања, изложени су релевантни постојећи резултати из проучаване области, нови резултати су јасно приказани и њихово тумачење је коректно. Обиман списак библиографских референци садржи релевантне радове и сведочи о добром познавању области. Дисертација је прегледна и добро организована.

3. По чему је дисертација оригиналан допринос науци

У оригиналне доприносе дисертације спадају:

1. Детаљна емпиријска анализа и тумачење утицаја *Sakoe-Chiba* ограничења на најчешће коришћене еластичне мере сличности у раду са временским серијама.
2. Темељно испитивање улоге глобалних ограничења у побољшању тачности класификације INN класификатора и предлог избора оптималне ширине појаса ограничења.
3. Испитивање важности најближег суседа при класификацији временских серија са  $k$ NN класификатором односно коришћења тежинске функције и глобалних ограничења за побољшање квалитета и стабилности  $k$ NN класификатора.
4. Проучавање могућности побољшања тачности класификације  $k$ NN класификатора применом различитих начина рачунања тежина и предлог најбољих решења у области временских серија.
5. Провера става да је тачност класификације једноставног INN класификатора тешко надмашити.
6. Дизајн и имплементација бесплатне и прошириве библиотеке отвореног кода која интегрише главне технике и методе потребне за анализу *i mining* временских серија.

4. Недостаци дисертације и њихов утицај на резултат истраживања

Дисертација нема недостатака.

**X ПРЕДЛОГ:**

На основу укупне оцене дисертације, комисија предлаже да се докторска дисертација под називом „Role of Similarity Measures in Time Series Analysis“ (Улога мера сличности у анализи временских серија) кандидата Золтана Гелера прихвати, а кандидату одобри јавна одбрана дисертације.

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

---

др Милош Радовановић, доцент, председник

---

др Мирјана Ивановић, редовни професор, ментор

---

др Зоран Будимац, редовни професор, члан

---

др Владимир Курбалија, ванредни професор, члан

---

др Зоран Боснић, ванредни професор, члан