

## НАСТАВНО-НАУЧНОМ ВЕЋУ

**Предмет:** Реферат о урађеној докторској дисертацији кандидата Жељка Недељковића.

Одлуком Наставно-научног већа Електротехничког факултета бр. 5003/13-3 од 14.10.2020. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Жељка Недељковића под насловом "Оптимизација система за аутоматско препознавање говорних емоција". 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 Жељко Недељковић је 23.01.2020. године пријавио тему за израду докторске дисертације под називом „Оптимизација система за аутоматско препознавање говорних емоција“.

Комисија за студије трећег степена размотрила је 03.03.2020. године предлог теме за израду докторске дисертације и предлог Комисије о оцени подобности теме и кандидата и упутила Наставно-научном већу Електротехничког факултета на усвајање.

Наставно-научно веће Електротехничког факултета је на својој седници од 10.03.2020. године, одлуком бр. 5003/13-1, именovalo Комисију за оцену услова и прихватање теме докторске дисертације у саставу др Александра Марјановић, доцент (Универзитет у Београду – Електротехнички факултет), др Зоран Перић, редовни професор (Универзитет у Нишу – Електронски факултет), др Татјана Лутовац, редовни професор (Универзитет у Београду – Електротехнички факултет) и др Драгана Шумарац Павловић, редовни професор, (Универзитет у Београду – Електротехнички факултет).

Дана 28.05.2020. године кандидат је полагао јавну усмену одбрану теме пред наведеном Комисијом и добио задовољавајућу оцену.

Наставно-научно веће Електротехничког факултета је на седници од 23.06.2020. године усвојило Извештај Комисије за оцену услова и прихватање теме докторске дисертације (Одлука бр. 5003/13-2).

Веће научних области техничких наука Универзитета у Београду дало је сагласност на предлог теме докторске дисертације (број 61206-2109/2-20 од 06.07.2020. године).

Дана 01.09.2020. године кандидат је предао докторску дисертацију на преглед и оцену.

Дана 06.10.2020. године Комисија за студије трећег степена потврдила је испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације.

На својој седници од 14.10.2020. године Наставно-научно веће Факултета именовало је Комисију за преглед и оцену докторске дисертације у саставу др Жељко Ђуровић, редовни професор (Универзитет у Београду – Електротехнички факултет), др Александра Марјановић, доцент (Универзитет у Београду – Електротехнички факултет), др Зоран Перић, редовни професор (Универзитет у Нишу – Електронски факултет) и др Драгана Шумарац Павловић, редовни професор (Универзитет у Београду – Електротехнички факултет) (број одлуке 5003/13-3).

На основу члана 101. Статута Универзитета у Београду, члана 74. Статута Универзитета у Београду-Електротехничког факултета и захтева студента, одобрено је продужење рока за завршетак студија до истека троструког броја школских година потребних за реализацију уписаног студијског програма.

## 1.2. Научна област дисертације

Докторска дисертација Жељка Недељковића под насловом „Оптимизација система за аутоматско препознавање говорних емоција“ припада научној области електротехника и рачунарство, а ужој научној области Обрада сигнала и управљање системима, за коју је матичан Електротехнички факултет Универзитета у Београду.

Ментор докторске дисертације је др Жељко Ђуровић, редовни професор Електротехничког факултета Универзитета у Београду. Проф. др Жељко Ђуровић се дуги низ година бави научноистраживачким радом у области обраде сигнала, стохастичких система и естимације као и примене различитих техника вештачке интелигенције, што је документовано релевантним научним радовима који су наведени приликом пријаве теме докторске дисертације кандидата.

## 1.3. Биографски подаци о кандидату

Жељко Б. Недељковић је рођен 10.06.1988. године у Ужицу. Гимназију је завршио са одличним успехом. Електротехнички факултет Универзитета у Београду је уписао 2007. године. Дипломирао је на модулу Сигнали и системи 2011. године са укупном просечном оценом 9,5. Мастер студије на Електротехничком факултету Универзитета у Београду је уписао 2011. године на модулу Сигнали и системи. Мастер студије је завршио са просечном оценом 10,0. Докторске студије на Електротехничком факултету Универзитета у Београду је уписао 2013. године на смеру Управљање процесима и обрада сигнала. Положио је све испите са просечном оценом 10,0. У току 2011. године је провео шест месеци на стручној пракси у фирми „Унит инжењеринг“ бавећи се развојем наменских рачунарских система. Од 2011. године је радно ангажован у ХЕ „Бајина Башта“ на пословима регулације, управљања и СКАДА система. Тренутно је на позицији главног инжењера за систем управљања и телекомуникације.

Област истраживања кандидата укључује обраду сигнала, препознавање облика, моделовање и индентификацију система, а ужа област истраживачког интересовања јесте карактеризација говорног сигнала и класификациони алгоритми. У самом фокусу његовог

рада се налазе проблеми везани за препознавање емоција из говорног сигнала, индентификацију говорника и препознавање говора. Истраживачким радом кандидата је руководио професор др Жељко Ђуровић. Досадашњи резултати су приказани кроз два рада у међународним часописима, пет радова на међународним конференцијама и један рад на националној конференцији.

## **2. ОПИС ДИСЕРТАЦИЈЕ**

### 2.1. Садржај дисертације

Докторска дисертација кандидата Жељка Недељковића под насловом „Оптимизација система за аутоматско препознавање говорних емоција“ је написана на 101 страни. Организована је у осам поглавља и садржи 30 слика и 15 табела. Налови поглавља су: Увод, Говор и емоције, Обележја говорног сигнала, Технике класификације, Анализа система за препознавање емоција у говору, Оптимизација система за препознавање емоција у говору, Закључак и Литература. У наставку дисертације је приказана кратка биографија кандидата, и на крају се налазе прилози. У даљем тексту је изложен кратак садржај појединих поглавља.

### 2.2. Кратак приказ појединачних поглавља

Докторска дисертација кандидата Жељка Недељковића састоји се од уводног поглавља, три прегледна поглавља, затим два истраживачка поглавља, праћена закључцима и прегледом литературе.

У уводној глави кандидат Жељко Недељковић говори о актуелности и сложености проблема аутоматског препознавања емоционалних стања говорника. Неизоставни део вештачке интелигенције јесте емоционална интелигенција, с обзиром да представља суштину највишег ступња интелигенције. Емоционална интелигенција почива на поузданом препознавању емоционалних стања, где се као најзначајнији извор емоционалних информација издваја паралингвистички садржај говора. Садашњи ниво научних достигнућа се може похвалити великим помацима на пољу аутоматског препознавања емоција из говора. Посматрајући хронолошки развој научне мисли у овој области, почев од првих експеримената у прошлом веку, па све до данас, могуће је створити јасну слику о технолошком развоју, али несумњиво се намеће мисао да и даље постоји простор за даља унапређења. Остварени успеси су већ нашли бројне примене у практичним апликацијама, при чему се издвајају примене у корисничким сервисима, едукацији, терапеутским третманима, безбедној вожњи, видео играма, као и у сврхе поправке квалитета система за аутоматско препознавање говора и говорника. На тај начин је уводно поглавље отворило простор за даље излагање о садржају и резултатима истраживања којима се кандидат бавио, али је истовремено посведочило о актуелности и комплексности проблема препознавања емоционалних стања говорника.

У другој глави је описан процес настајања говора и представљен појам емоција, а потом је описана манифестација емоција путем говора. Говор има врло битну улогу у процесу комуникације међу људима, при чему је доминантна лингвистичка порука, али је неизоставна и нелингвистичка компонента (пол, старост, емоционално стање,...), која садржи вредне информације и уобличава основу поруку. Лингвистичка компонента поруке може, али не мора, садржати информације о емоционалном стању говорника, или може

садржати информације које нису тачне. Зато се нелингвистичка компонента сматра вреднијом са аспекта препознавања емоционалних стања говорника. Током емоционалног доживљаја долази до низа промена у организму јединке, што за последицу има одступање одређених функција у односу на неутрално стање, тако да се могу препознати обрасци промена акустичких параметара специфични за различита емоционална стања, а који се јављају као последица утицаја физиолошких промена на дисање, изговор и артикулацију. Дефинисање броја емоционалних стања и границе између њих је врло важно са аспекта њихове класификације, тако да су у наставку поглавља приказане основне идеје категоричких и димензионалних емоционалних модела.

У трећој глави је дат преглед обележја која могу бити употребљена за репрезентацију говорног сигнала. Неоспорно је да су у снимљеном говорном сигналу садржане информације о емоционалном стању говорнику, при чему је неизоставни део и лингвистичка порука, као и информације својствене за говорника, попут боје гласа, пола, старости, али и карактеристике преносног канала. Класификација емоционалних стања употребом неизмењеног говорног сигнала не би била ефикасна, стога је потребно извршити издвајање обележја говорног сигнала која што верније осликавају емоционална стања уз потискивање осталих садржаја. У овој глави су дате теоријске основе обележја говора из групе спектралних обележја као што су *LPCC* (енг. *Linear Prediction Cepstral Coefficients*), *MFCC* (енг. *Mel Frequency Cepstral Coefficients*), *LFPC* (енг. *Log Frequency Power Coefficients*), континуалних обележја (фундаментална фреквенција, енергија, форманти, обележја трајања), обележја квалитета гласа (цитер (енг. *jitter*), шимер (енг. *shimmer*), *HNR* (енг. *Harmonic to Noise Ratio*), *ZCR* (енг. *Zero Crossing Rate*), вероватноћа звучног говора)) и нелинеарних обележја (Тајгер (енг. *Teager*) енергија). Већина наведених обележја врше кратковремено описивање говорног сигнала, а како би се дошло до компактне репрезентације комплетног говорног исказа у употреби су различите статистике говорног сигнала. Употребом већег броја обележја се укључује и већа количина релевантних информација за репрезентацију емоционалних стања, а са друге стране се повећава димензионалност вектора обележја, тако да су размотрене и технике за редукцију димензија вектора обележја. У том смислу је посебна пажња посвећена линеарној дискриминантној анализи (*LDA* – енг. *Linear Discriminant Analysis*).

Четврта глава се бави класификационим техникама које имају за задатак да донесу одлуку о садржаном емоционалном стању у датом исказу на основу издвојених обележја. Детаљније су описане технике коришћене у раду: *HMM* (скривени Марковљеви модели, енг. *Hidden Markov Models*) као традиционалан метод у обради говора, *SVM* (метода потпорних вектора, енг. *Support Vector Machines*) који се показао супериорним у многим задацима препознавања облика и *DNN* (дубоке неуралне мреже, енг. *Deep Neural Networks*) као репрезент дубоког учења, приступа који је у повоју у обради говора. Од мноштва доступних класификаторских метода препознају се предности различитих приступа, тако да се адекватним комбиновањем различитих приступа могу постићи боље перформансе у односу на индивидуалне примене. С тим у вези, у овој глави су дате и основе мултикласификаторских система. Након што се формира класификаторска структура, изузетно је важно проценити перформансе пројектованог система, тако да су представљене и мере квалитета класификаторских система.

Пета глава је посвећена анализи квалитета обележја и класификатора означеним као добар избор за потребе аутоматског препознавања говорних емоција. Уједно је извршена анализа утицаја коришћених база емоционалног говора на резултате тестирања. У анализи су коришћена спектрална обележја: *LPCC*, *LFPC* и *MFCC*. Од класификатора су укључени: *HMM*, *SVM* и *DNN*. Тестови су спроведени на базама српског, руског, пољског и немачког емоционалног говора. Понављањем експеримената за све комбинације обележја,

класификатора и база у поставци независно од говорника и поставци зависно од говорника је продуковало велики број резултата. Како би се олакшало извођење закључака, придружена је мера успешности сваком обележју, сваком класификатору и сваком пару обележје-класификатор. На основу тако уведених мера је било могуће изоловати перформансе појединачних елемената система, као и сагледавање утицаја база на перформансе система.

У шестој глави је акценат био на унапређењу система за аутоматско препознавање говорних емоција, ослањајући се при томе на резултате компаративне анализе из претходне главе. Првенствено су анализирани могућности унапређења класификационе технике на бази дискретних скривених Марковљевих модела, при чему се векторска квантизација истиче као јако изазован задатак, тако да је њој посвећена посебна пажња. Размотрено је неколико приступа заснованих на употреби *QQ* криве (енг. *Quantile-Quantile plot*) код одређивања сталона за векторску квантизацију. Након што је усвојена коначна форма векторске квантизације, размотрене су могућности додатног унапређења *HMM*. Уведен је корак предобуке модела, као и вишеструка обука са комбиновањем модела. Предложена *HMM* имплементација је тестирана употребом обележја и база са којима је тестирана и стандардна *HMM* имплементација, тако да су остварени резултати недвосмислено потврдили да предложена унапређења доприносе поправци перформанси у односу на стандардну имплементацију. У претходним анализама су превасходно коришћена спектрална обележја говорног сигнала, с обзиром на опште уверење да су добар избор за задатак аутоматског препознавања говорних емоција, што не искључује могућност постојања корисних информација и у другим групама обележја. Верификација постојања корисних информација изван спектралних обележја је извршена тестирањем скупа обележја сачињеног од представника преосталих група. С обзиром да су неспектрална обележја постигла уједначене перформансе са спектралним обележјима, и с обзиром на различите природе обележја, закључено је да се њиховим комбиновањем могу обезбедити потпуније информације. Укључењем великог броја представника различитих група обележја у јединствен проширени скуп обележја је обезбеђена значајна поправка перформанси. Са проширењем скупа обележја се јавља проблем манипулације векторима великих димензија, такође, долази до проређења простора обележја, тако да је било неопходно разматрање могућности редукације димензија вектора обележја, уз што мање губитке корисних информација. Експерименти су спроведени употребом *LDA* методе за редукацију димензија, а анализирана је и могућност редукације димензија употребом Фибоначијевог низа бројева. Након што су испитане могућности различитих поправки појединачних елемената класификаторских система, преостала је да се размотри интеграција предности различитих приступа у јединствен систем. Конструисана је паралелна мултикласификаторска структура са комбинаторним правилом које поред резултата класификације датог класификаторског система користи и информације о карактеристикама класификаторског система израженим кроз матрицу конфузије и тачност препознавања. Предложено комбинаторно правило је тестирано са класификаторским системима формираним приликом спровођења претходних експеримената, а поређењем резултата мултикласификаторске структуре са резултатима појединачних класификаторских система је потврђена поправка тачности препознавања. Такође, размотрено је формирање ансамбла класификаторских система произвољне величине употребом Фибоначи редукације димензија, што је за већину база допринело додатној поправци резултата.

Седмо поглавље садржи резиме тезе и сублимиране закључке. У овом поглављу је извршена кратка рекапитулација проблема који су били у фокусу истраживања, предложених решења, као и учених ограничења и смерница за будућа истраживања. На крају је дат и преглед коришћене литературе.

### 3. ОЦЕНА ДИСЕРТАЦИЈЕ

#### 3.1. Савременост и оригиналност

Данашњи развој друштва диктира све строжије захтеве према техничким системима, па се дошло дотле да је вештачка интелигенција као основни задатак добила да омогући машинама да обављају различите функције као што би их обављали људи. При томе се захтева и нешто више, а то је да постоји и интеракција која личи на хуману. Већ је много урађено на пољу интеракције човек-машина, где су већ присутни успешни системи за аутоматско препознавање природног говора као и за синтезу говора. Међутим, у свакој комуникацији су вредне информације садржане и у ономе што није експлицитно речено. Ту се долази до појма емоционалне интелигенције као суштине машина које су много општијег карактера, робусније, ефикасније, али и усклађеније са хуманим вредностима. Емоционална интелигенција подразумева способност разумевања осећања и расположења, као и управљање понашањем у складу са њима, што у рачунарским оквирима подразумева постојање алгоритама за препознавање емоционалних стања, алгоритама за разматрање емоција у доношењу одлука, као и алгоритама за симулацију емоционалних стања. С обзиром да је поуздано препознавање емоција основ за правилан рад система који се баве осталим аспектима емоционалне интелигенције, суштински је битно да препознавање буде извршено са високим степеном поузданости. Информација о емоционалном стању је кодирана у говору, изразима лица, покретима тела и физиолошким знацима. При томе се паралингвистички садржај говора може сматрати најзначајнијим носиоцем емоционалних информација с обзиром на заступљеност вербалне комуникације, степен поузданости, као и могућност прикупљања.

Предмет истраживања кандидата јесу управо алгоритми за аутоматско препознавање емоција из паралингвистичког садржаја говора. Направљени су велики помаци на пољу аутоматског препознавања емоција из говора од првих експеримената до данас, али свакако постоји простор за даља унапређења. У ранијим истраживањима су коришћени различити приступи, где се у сваком може разликовати део који се бави издвајањем обележја говорног сигнала и део који се бави класификацијом издвојених обележја. У раду кандидата је спроведена компаративна анализа обележја и класификатора најчешће коришћених у решавању задатка аутоматског препознавања емоционалних стања говорника, при чему су за поређење перформанси елемената система коришћени показатељи које је кандидат предложио. Потом су размотрена различита унапређења која су подразумевала употребу потпуно нове методе векторске квантизације код дискретних *НММ* модела, испитивање подобности различитих репрезентата говорног сигнала, ангажовање оригиналне методе за редукцију димензија вектора обележја засноване на Фибоначијевом низу бројева, конструкцију паралелне мултикласификаторске структуре са оригиналним предлогом комбинаторног правила, као и предлог аутоматизованог формирања ансамбла класификатора произвољних димензија применом Фибоначи редукције димензија.

По сазнању аутора тезе и чланова Комисије, у литератури тренутно не постоје методе које наведене проблеме адресирају на овако својствен начин.

#### 3.2. Осврт на референтну и коришћену литературу

Списак коришћене литературе броји 127 наслова и она се, грубо говорећи, може поделити у три скупа. Први од њих представља базичну литературу из које се могу сагледати основе

техника за репрезентацију говорног сигнала, класификационих и других техника машинског учења примењивих у овој области, као и структура система за аутоматско препознавање говорних емоција, али и проблеми који су скопчани са предметним задатком. Иако је број референци у овој области изузетно велики, свакако као њихове репрезенте треба поменути следеће наслове: *L. Rabiner, B.-H. Juang, Fundamentals of Speech Recognition, New Jersey: Prentice Hall, 1993.*; *T. L. Nwe, S. W. Foo, L. C. D. Silva, „Speech Emotion Recognition Using Hidden Markov Models,” Speech Communication, vol. 41, no. 4, pp. 603-623, 2003.*; *C. Cortes, V. Vapnik, „Support-Vector Networks,” Machine Learning, vol. 20, pp. 273-297, 1995.*; *L. Deng, D. Yu, „Deep Learning: Methods and Applications.” Foundations and Trends in Signal Processing, vol. 7, no. 3-4, pp. 197-387, 2013.*

Другу групу референтних наслова чине резултати објављени скоријих година и који сведоче како о актуелности проблема који се разматрају тако и о различитим приступима на који се они могу решавати. У овој групи коришћене и реферисане литературе свакако треба поменути следеће наслове: *D. Schuller, B. W. Schuller, „The Age of Artificial Emotional Intelligence,” Computer, vol. 51, no. 9, pp. 38-46, 2018.*; *B. W. Schuller, „Speech Emotion Recognition: Two Decades in a Nutshell, Benchmarks, and Ongoing Trends,” Communications of the ACM, vol. 61, no. 5, pp. 90-99, 2018.*; *Y. B. Singh, S. Goel, „Survey on Human Emotion Recognition: Speech Database, Features and Classification,” in 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN, Greater Noida, 2018.*; *M. M. E. Ayadi, M. S. Kamel, F. Karray, „Survey on Speech Emotion Recognition: Features, Classification Schemes, and Databases,” Pattern Recognition, vol. 44, no. 3, pp. 572-587, 2011.*

Трећу групу референтних радова чине управо резултати који се налазе у самој сржи ове докторске дисертације. Они документују истраживања и напоре да се нађу ефикасне методе за препознавање емоционалних стања употребом различитих база емоционалног говора. У овој групи наслова свакако треба поменути следеће: *V. Delić, M. Bojanić, M. Gnjatović, M. Sečujski, S. T. Jovičić, „Discrimination Capability of Prosodic and Spectral Features for Emotional Speech Recognition,” Elektronika ir Elektrotehnika, vol. 18, no. 9, pp. 51-54, 2012.*; *A. Hassan, R. I. Damper, „Multi-Class and Hierarchical SVMs for Emotion Recognition,” in Proceedings of the 11th Annual Conference of the International Speech Communication Association, Makuhari, 2010.*; *A. Shaukat, K. Chen, „Emotional State Recognition from Speech via Soft-Competition on Different Acoustic Representations,” in The 2011 International Joint Conference on Neural Networks, San Jose, 2011.*; *D. Kamińska, T. Sapiński, G. Anbarjafari, „Efficiency of Chosen Speech Descriptors in Relation to Emotion Recognition,” EURASIP Journal on Audio, Speech, and Music Processing, p. 2017:3, 2017.*; *P. Staroniewicz, „Automatic Recognition of Emotional State in Polish Speech,” in Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues, Lecture Notes in Computer Science, Berlin, 2011.*

Увидом у наведену литературу може се стећи утисак о томе да је, радећи на својој докторској дисертацији, кандидат користио релевантну расположиву научну литературу која је у директној вези са проблемом који је обрађиван у дисертацији. На овај начин је кандидат уверио чланове Комисије да резултати изложени у овој тези заиста представљају помак у научном и истраживачком опусу који се бави третираним проблемом. Са друге стране, на овај начин се стиче увид у чињеницу да је своје резултате кандидат поредио са најновијим светским резултатима из дотичне области. Такав приступ омогућава научној јавности да на непосредан и једноставан начин формира објективан суд о вредности садржаја и резултата који су у докторској дисертацији изложени.

### 3.3. Опис и адекватност примењених научних метода

Током израде своје докторске дисертације кандидат Жељко Недељковић је од научних метода истраживања доминантно користио технике анализе, синтезе, компарације, генерализације, експеримента и статистике.

С почетка, улазећи у научну област обраде говорног сигнала и специфично, аутоматског препознавања говорних емоција, кандидат је био посвећен евалуацији квалитета индивидуалних елемената система за аутоматско препознавање говорних емоција. Овај сегмент истраживања се доминантно ослонио на научни истраживачки метод анализе. Квалитет обележја и класификатора коришћених у конструкцији система за аутоматско препознавање говорних емоција је евалуиран понављањем тестова за различите комбинације обележја, класификатора и база, а затим одређивањем мера квалитета обележја, класификатора и њихових парова, тако да су овај сегмент истраживања карактерисале и научне методе компарације, експеримента и статистике.

Касније, улазећи дубље у област истраживања, кандидат је предложио поступак векторске квантизације код дискретних *HMM* који подразумева употребу Фибоначијевог низа бројева за потребе трансформације оригиналних вредности и њиховом груписању применом *QQ* криве, где се препознаје доминантна употреба научних истраживачких метода синтезе и генерализације. Истраживачка метода синтезе се доминантно уочава и у делу истраживања које се бави фузијом већег броја класификаторских структура у здружену мултикласификаторску структуру. Истраживањем је обухваћено и испитивање репрезентације говорног сигнала употребом различитих обележја, а сва тестирања су вршена поступним увођењем нових метода, уз одређивање мере квалитета предложених проширења, тако да је и у овом сегменту истраживања могуће препознати употребу научних истраживачких техника експеримента и статистике. Сва тестирања су вршена употребом расположивих база, изводећи при томе опште закључке о ваљаности предложених метода, што имплицира примену научне истраживачке методе генерализације.

### 3.4. Применљивост остварених резултата

Резултати остварени током рада на овој докторској дисертацији имају вишеструку применљивост. Закључци компаративне анализе типичних елемената система за аутоматско препознавање говорних емоција могу помоћи код избора најбоље комбинације обележја и класификатора приликом пројектовања система за аутоматско препознавање говорних емоција, али и послужити као основа за будућа истраживања.

Предложена метода за векторску квантизацију код дискретних *HMM* заснована на *QQ* кривој премашује резултате постигнуте применом уобичајене методе засноване на *K-means* кластеризацији. Такође, предложена метода може наћи примену у решавању произвољног кластеризационог проблема који представља незаобилазни сегмент готово сваког система за препознавање облика или система за асистенцију при доношењу одлука.

Предложена је и метода за редукцију димензија вектора обележја заснована на примени Фибоначијевог низа бројева, коју као основно својство карактерише једноставност и рачунарска ефикасност. Предложена метода за редукцију димензија је у спроведеним тестовима успела да надмаши афирмисану *LDA* методу редукције димензија. Потенцијал предложене методе може доћи до изражаја у применама са акцентом на рачунској



ефикасности и на местима где нису познате, или су сувише сложене, зависности задатих података.

На крају, у раду је предложено комбинаторно правило за формирање здруженог резултата ансамбла класификатора паралелне мултикласификаторске структуре. Предложено је и аутоматско формирање ансамбла класификатора произвољних димензија применом Фибоначи редукције димензија, без потребе за упуштање у природу проблема који се решава.

### 3.5. Оцена достигнутих способности кандидата за самостални научни рад

Током рада на својој докторској дисертацији кандидат је показао да је у стању да самостално уочава проблеме, да из доступне литературе препознаје елементе који су од веће или мање важности за њихово решавање. Такође је, што је нарочито важно, стекао могућност да критички анализира постојећа решења. Током времена, напредујући кроз рад, испољио је захтевану научну зрелост и оспособљен је за даљи успешан научно-истраживачки рад. Ова је научна одмереност нарочито дошла до изражаја у сегментима оцене вредности постигнутих научних резултата и у процесу уочавања ограничења и мањкавости сопствених развијених метода за кластеризацију вишедимензионих вектора, редукцију димензија и формирање мултикласификаторске структуре.

## **4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС**

### 4.1. Приказ остварених научних доприноса

Током истраживања и рада на докторској дисертацији кандидат је остварио врло значајне резултате који се могу исказати кроз следеће научне доприносе:

- Извршен је врло детаљан преглед научне области која се бави препознавањем емоционалних стања из говорног сигнала моделовањем говорног сигнала и техникама класификације. На критичан начин су наведена ограничења и недостаци појединих приступа и на тај начин је отворен простор за развој нових метода који су описане у овој дисертацији.
- Извршена је компаративна анализа спектралних обележја говорног сигнала према степену подобности за употребу у системима аутоматског препознавања говорних емоција.
- Такође је на основу исцрпне статистичке анализе, а коришћењем база емотивног говора за различите европске језике, извршено поређење типичних класификационих алгоритама у односу на применљивост у системима аутоматског препознавања говорних емоција.
- Постављене су референтне вредности у смислу вероватноћа исправног препознавања за најчешће коришћене поставке обележје-класификатор за словенске базе емоционалног говора.
- Извршен је предлог мера квалитета које омогућавају изоловање перформанси појединачних елемената система за аутоматско препознавање говорних емоција и сагледавање утицаја база на перформансе система.

- Предложена је нова метода за векторску квантизацију код дискретних скривених Марковљевих модела, као и друга унапређења која повољно утичу на квалитет формираног модела.
- Предложена је нова метода за редукцију димензија вектора обележја заснована на Фибоначијевом низу бројева.
- Предложено је комбинаторно правило паралелне мултикласификаторске структуре, као и поступак за формирање ансамбла класификатора произвољне величине.

#### 4.2. Критичка анализа резултата истраживања

У жижи истраживања које је приказано у оквиру ове докторске дисертације је било унапређење система за аутоматско препознавање говорних емоција, што је обухватило увођење нове методе за векторску квантизацију код дискретних *НММ*, нове методе за редукцију димензија, као и предлог паралелне мултикласификаторске структуре. Предложена унапређења су омогућила достизање циљане поправке перформансе система за аутоматско препознавање говорних емоција, а уједно представљају универзална решења предметних задатака која могу наћи примену као таква у конструкцији других система за аутоматско препознавање говорних емоција. Истовремено ови резултати и њихова примена не морају нужно бити лимитирани на препознавање емоција из говора већ могу наћи примену у решавању ширег скупа задатака. Предложене методе унапређују перформансе система за аутоматско препознавање говорних емоција, али не омогућавају тотално елиминисање грешака у препознавању емоционалних стања, што оставља простор за будућа истраживања. Такође, предложена унапређења су тестирана на базама одглумљеног емоционалног говора, што је створило повољне услове за изражајно осликавање карактеристика различитих емоционалних стања приликом развоја нових алгоритама, али за евалуацију пројектованог система у условима блиским реалној примени је прикладнија примена база спонтаног емоционалног говора, што може бити тема будућих истраживања.

На крају, Комисија изражава задовољство што су сви очекивани научни доприноси, наведени приликом пријаве теме истраживања, остварени и детаљно презентовани у коначној верзији докторске дисертације. Такође, Комисија је задовољна формом и начином презентовања остварених резултата.

#### 4.3. Верификација научних доприноса

У циљу верификације наведених научних доприноса, наводимо списак радова које је кандидат Жељко Недељковић публиковао радећи на својој докторској дисертацији:

##### Категорија М23:

1. **Nedeljković Ž.**, Milošević M., Đurović Ž.: Analysis of features and classifiers in emotion recognition systems: Case study of Slavic languages, - *Archives of Acoustics*, vol. 45, no. 1, pp. 129–140, 2019 (DOI: 10.24425/aoa.2020.132489).
2. Milošević M., **Nedeljković Ž.**, Glavitsch U., Đurović Ž.: Speaker Modeling Using Emotional Speech for More Robust Speaker Identification, - *Journal of Communications Technology and Electronics*, vol. 64, no. 11, pp. 1256-1265, 2019 (DOI: 10.1134/S1064226919110184).

#### Категорија М33:

1. **Nedeljković Ž.**, Milošević M., Đurović Ž.: Improved HMM for automatic speech emotion recognition, - *Proceedings of XXVII Pacific Voice Conference*, Krakow, Poland, 2020, vol. 3, pp. 10-14.
2. **Nedeljković Ž.**, Đurović Ž.: QQ-plot based clustering, - *Proceedings of 6th International Conference on Electrical, Electronic and Computing Engineering*, Silver Lake, Serbia, 2019, pp. 138-141.
3. **Nedeljković Ž.**, Milošević M., Đurović Ž.: Baseline system for speaker recognition – parameter analysis, - *Proceedings of 4th International Conference on Electrical, Electronics and Computing Engineering*, Kladovo, Serbia, 2017, pp. AUII.3.1-5.
4. Milošević M., **Nedeljković Ž.**, Đurović Ž.: Emotional speech databases in Slavic languages – an overview, - *26th Annual Conference of the International Association for Forensic Phonetics and Acoustics*, Split, Croatia, 2017, pp. 167-170.
5. Milošević M., **Nedeljković Ž.**, Đurović Ž.: SVM classifier for emotional speech recognition in software environment SEBAS, - *Proceedings of 3rd International Conference on Electrical, Electronic and Computing Engineering*, Zlatibor, Serbia, 2016, pp. AUI4.1.1-4.

#### Категорија М63:

1. **Nedeljković Ž.**, Đurović Ž.: Automatsko prepoznavanje emocija na osnovu govora upotrebom skrivenih Markovljevih modela, - *Zbornik 59. konferencije za elektroniku, telekomunikacije, računarstvo, automatiku i nuklearnu tehniku ETRAN 2015*, Srebrno jezero, Srbija, 2015, str. AUI.6.1-5.

## 5. ЗАКЉУЧАК И ПРЕДЛОГ

Докторска дисертација кандидата Жељка Недељковића под називом „Оптимизација система за аутоматско препознавање говорних емоција“ је у целини написана у складу са образложењем наведеним у пријави теме и садржи све битне елементе који се захтевају Правилником о докторским студијама Електротехничког факултета Универзитета у Београду.

У дисертацији се на систематски и доследан начин анализира проблем поправке перформанси система за аутоматско препознавање емоција из говора, приликом чега је предложено унапређење дискретних скривених Марковљевих модела које подразумева ангажовање потпуно нове методе векторске квантизације. Поред тога је предложена нова метода за редукцију димензија вектора обележја, као и комбинаторно правило за конструкцију паралелне мултикласификаторске структуре са предлогом за формирање ансамбла класификатора произвољних димензија. Предложена унапређења успевају да поправе перформансе система за аутоматско препознавање емоција, при чему су предложене методе примењиве и на много шири скуп проблема.

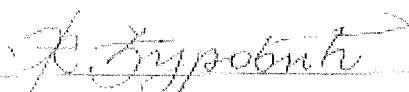
Кандидат Жељко Недељковић је показао способност за самостални научни рад, што потврђује и чињеница је објавио неколико научних радова који су проистекли из истраживања приказаних у оквиру ове тезе, а у којима се појављује као први коаутор. Оцењујући докторску дисертацију, као и чињеницу да је анализирани и решавани проблем актуелан и савремен, као и да садржи научне доприносе, Комисија констатује да је

кандидат Жељко Недељковић, мастер инжењер електротехнике и рачунарства, испунио све услове предвиђене Законом о високом образовању, Статутом и Правилником о докторским студијама Електротехничког факултета Универзитета у Београду.

Имајући у виду напред наведено, Комисија са задовољством предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да се докторска дисертација под називом „Оптимизација система за аутоматско препознавање говорних емоција“ кандидата Жељка Недељковића прихвати, изложи на увид јавности и упути на коначно усвајање Већу научних области техничких наука Универзитета у Београду

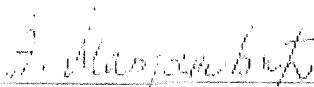
У Београду, 19.11.2020.

#### ЧЛАНОВИ КОМИСИЈЕ



др Жељко Ђуровић, редовни професор (ментор)

Универзитет у Београду - Електротехнички факултет



др Александра Марјановић, доцент

Универзитет у Београду - Електротехнички факултет



др Зоран Перић, редовни професор

Универзитет у Нишу - Електротехнички факултет



др Драгана Шумарац Паčkинић, редовни професор

Универзитет у Београду - Електротехнички факултет