

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 03.09.2024. године именовала нас је у Комисију за преглед и оцену мастер рада дипл. инж. Сања Дрча под насловом „Примена Витербијевог алгоритма за препознавање карактера“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Сања Дрча је рођена 10.01.1998. године у Београду. Завршила је основну школу "Ратко Митровић" у Београду, а затим Девету гимназију „Михаило Петровић Алас“, такође, у Београду. Електротехнички факултет је уписала 2016. године. Дипломирала је на одсеку Телекомуникације и информационе технологије са просечном оценом 8.17. Дипломски рад на тему „Имплементација функционалности динамичког веб сајта“ одбранила је у септембру 2020. године са оценом 10. Мастер академске студије на Електротехничком факултету у Београду, на модулу Информационо комуникационе технологије уписала је у октобру 2020. године. Положила је све испите са просечном оценом 9.8.

2. Извештај о студијском истраживачком раду

Кандидат Сања Дрча је као припрему за израду мастер рада урадила преглед релевантне литературе која се односи на препознавање образца и технике које се користе за то, а потом се определила да у свом мастер раду детаљније опише и истражи оптичко препознавање карактера (*Optical Character Recognition*) уз употребу Витербијевог алгоритма. Предмет рада представља опис поступка за реконструкцију образца на примеру текста који може имати униформну или насумичну расподелу слова. Наиме, претпоставља се да се сваки карактер текста може представити матричним записом, тако да сваки члан матрице има вредност 0 или 1. На насумичан начин су изабрани неки чланови матрица којима су инвертоване вредности тако да ће на неким позицијама на којима је била јединица, бити уписана нула, и обрнуто. Главни циљ рада је да се од новонасталих матрица са неким инвертованим вредностима препозна како је изгледала оригинална матрица и очита који карактер је она представљала. У те сврхе се користи Витербијев алгоритам, а код за то је кандидат написао у програмском језику *Matlab*.

3. Опис мастер рада

Мастер рад обухвата 45 страна, са укупно 34 слике и 9 табела. Рад садржи увод, 4 поглавља и закључак (укупно 6 поглавља), списак скраћеница, списак слика, списак табела и списак коришћене литературе.

Прво поглавље представља увод у коме су описаны предмет и циљ рада.

У другом поглављу дат је шематски приказ телекомуникационог система са описом свих блокова који га чине.

Треће поглавље је посвећено опису скривених Марковљевих ланаца и Витербијевом алгоритму који се примењује за реконструкцију изобличених образца.

У четвртом поглављу је описан процес оптичког препознавања карактера кроз осам корака: оптичко скенирање, сегментација локације, предоброда, сегментација, репрезентација, екстракција карактеристика, обука и препознавање карактера и пост-оброда. Такође, дат је пример који пролази кроз све фазе оптичког препознавања карактера уз употребу Витербијевог алгоритма.

У петом поглављу је дат конкретан пример примене Витербијевог алгоритма на реконструкцију два текста, од којих један има унiformну расподелу слова, а други насумичну. Први текст се састоји од 6080 карактери, који могу бити неко од слова српске ћирилице, као и размак и тачка. Други текст, који има насумичну расподелу слова, састоји се од 6919 карактера. Свако слово српске ћирилице је представљено матрицом димензија 5x5, 8x8 и 10x10. Приликом учитавања ових текстова, може доћи до одређених грешака, тако да нека слова могу бити нејасна, односно њихове матрице се могу разликовати од оригиналних којима су првобитно представљене. Циљ овог рада је да покаже како се ради реконструкција изобличеног текста употребом Витербијевог алгоритма који за проналажење изворног текста користи метрику Хеминговог растојања. Резултати добијени коришћењем матрица различитих димензија су упоређени, а код за то је написан у програмском језику *Matlab*.

Шесто поглавље је закључак у оквиру кога се резимирају резултати рада. Потом је дат списак коришћених референци, списак скраћеница, списак слика и списак табела.

4. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Сање Дрча се бави анализом поступка и техника за оптичко препознавање карактера на примеру два текста, од којих један има унiformну расподелу слова, а други насумичну. Сваки карактер који се налази у та два текста може бити неко од слова српске ћирилице, као и размак и тачка и представљен је матрицама димензија 5x5, 8x8 и 10x10. Приликом скенирања текста долази до одређених грешака на појединим словима тако да се матрице којима су представљена разликују од оригиналних. Циљ рада јесте да се изврши реконструкција скенираног текста тако да се препозна шта је извршено било написано. Препознавање изворних карактера из изобличених се ради употребом Витербијев алгоритам који користи метрику Хеминговог растојања. Поступак препознавања карактера се понавља по три пута за оба текста. Први пут су сви карактери представљени матрицама димензијама 5x5, други пут са 8x8 и трећи пут су коришћене матрице димензија 10x10 за свако слово.

Резултат ће показати да се употребом матрица већих димензија добија боља прецизност реконструкције, односно мања су одступања изобличеног текста од изврног тако да се оне користе у системима где је тачност веома битна, као што је случај у медицинским системима. Међутим, главна мана употреба матрица већих димензија јесте та што је потребно доста времена да се изврши цео процес реконструкције текста, тако да ако је циљ брзина извршавања, а одређени ниво грешака прихватљив, матрице мањих димензија су погоднији избор.

5. Закључак и предлог

Кандидат Сања Дрча је у свом мастер раду успешно описала све фазе које се користе приликом оптичког препознавања карактера, са посебним акцентом на најважнијем кораку самог препознавања који је одрађен у програмском језику *Matlab*. Добијене резултате за реконструкцију оригиналног текста из изобличеног када се користе матрице различитих димензија за представљање сваког карактера, приказала је графички и упоредила. Поред тога, описала је главне предности и недостатке приликом употреба матрица већих и мањих димензија. Сања је показала да добро влада материјом коју је описала у раду и исказала самосталност у писању кода којим је покрила конкретне примере реконструисања текстова.

На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Сање Дрча прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 12.09.2024. године

Чланови комисије:

Др Предраг Иваниш, редовни професор

Др Весна Благојевић, ванредни професор