

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 14.10.2025 године именовала нас је у Комисију за преглед и оцену мастер рада дипл. инж. Урош Козић под насловом „Анализа географске пристрасности модела за препознавање именованих ентитета у српском језику“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Урош Козић је рођен 11.01.2002. године у Београду. Завршио је основну школу „Ђорђе Крстић“ у Београду као вуковац. Уписао је Трећу београдску гимназију, коју је завршио са одличним успехом. Електротехнички факултет уписао је 2020. године. Дипломирао је у року на одсеку за Софтверско инжењерство. Дипломски рад одбранио је у септембру 2020. године са оценом 10. Дипломске академске – мастер студије на Електротехничком факултету у Београду, на Модулу за Софтверско инжењерство уписао је у октобру 2024. године. Тренутно је запослен у компанији Ерат као софтверски инжењер.

2. Извештај о студијском истраживачком раду

Кандидат Урош Козић је као припрему за израду мастер рада урадио истраживање релевантне литературе из области препознавања именованих ентитета, пристрасности модела за обраду природног језика и географске пристрасности. Посебно су анализирани радови који се баве родном и географском пристрасношћу NER и геопarsing система, као и карактеристике модела за српски и сродне јужнословенске језике, укључујући CLASSLA, BERTiс и ICEF COMtext. Анализом постојећих истраживања утврђено је да је географска пристрасност NER модела за српски језик недовољно истражена, што је послужило као основа за дефинисање предмета, хипотеза и експерименталне поставке мастер рада.

3. Опис мастер рада

Мастер рад обухвата 86 страна рачунајући од увода, од чега прилози обухватају 3 стране, са укупно 26 слика, 14 табела и 14 референци. Рад садржи увод, четири тематска поглавља и закључак (укупно 6 поглавља), списак коришћене литературе, прилоге, списак скраћеница, слика и табела.

Прво поглавље представља увод у коме су описани предмет и циљ рада, значај испитивања географске пристрасности у системима за препознавање именованих ентитета и основне истраживачке хипотезе. Посебан фокус стављен је на препознавање адресних израза из Србије и Хрватске и на факторе који могу утицати на разлике у успешности модела.

У другом поглављу дат је преглед постојећих истраживања у области препознавања именованих ентитета и пристрасности у обради природног језика. Представљена су релевантна истраживања родне и географске пристрасности, као и радови и ресурси који се односе на NER моделе за српски и сродне јужнословенске језике.

У трећем поглављу описани су прикупљање и обрада података. Представљени су извори адресних података за Србију и Хрватску, поступак њиховог усклађивања, нормализације и чишћења, као и формирање скупа адресних израза и реченичних примера различитих домена. Размотрени су и основни изазови и ограничења формираног скупа података.

Четврто поглавље описује методологију истраживања. Представљени су модели CLASSLA, BERTiс и две варијанте ICEF COMtext модела, поступак уједначавања њихових излазних ознака, евалуационе метрике, формирање реченичних шаблона и практична имплементација експеримената са реченичним контекстом и без њега.

У петом поглављу приказани су и анализирани резултати. Испитана је географска пристрасност на нивоу Србије и Хрватске, утицај реченичног контекста, разлике између појединачних градова, као и утицај компоненти сложених адресних структура на успешност и географску стабилност модела.

Шесто поглавље представља закључак рада.

4. Анализа са кључним резултатима

Мастер рад дипл. инж. Уроша Козића бави се анализом географске пристрасности модела за препознавање именованих ентитета у српском језику, са посебним фокусом на адресне изразе из Србије и Хрватске. У

раду су испитана четири NER модела BERTić, CLASSLA, као и екавска и ијекавска варијанта ICEF COMtext модела, у више контролираних експерименталних поставки које обухватају самосталне адресне изразе, реченични контекст, поређења на нивоу градова и анализу компоненти сложених адреса.

За потребе истраживања формиран је обиман скуп адресних израза различите структурне сложености из Србије и Хрватске, уз контролисане реченичне примере из више текстуалних домена. Географска пристрасност анализирана је поређењем перформанси модела између држава и градова, док је додатном анализом испитан утицај реченичног и структурног контекста на успешност и стабилност препознавања појединачних компоненти адресе.

Основни резултати рада су: 1) потврђено постојање мерљивих географских разлика у перформансама NER модела, при чему правац и интензитет пристрасности зависе од конкретног модела и адресне компоненте; 2) показано је да реченични контекст повећава укупну успешност и на агрегатном нивоу смањује географске разлике, али их не уклања код свих структура; 3) утврђено је да значајне разлике постоје и између градова унутар исте државе и да исти градски пар може мењати релативну предност у зависности од типа адресног ентитета; 4) показано је да ефекат структурног контекста зависи од врсте додате информације, при чему кућни број има изражен стабилизујући ефекат на уличну компоненту, док богатији контекст највише доприноси стабилности назива насеља; 5) формиран је скуп података намењен анализи географске пристрасности NLP модела за српски језик, који може послужити као основа за даља истраживања.

5. Закључак и предлог

Кандидат Урош Козић је у свом мастер раду успешно спровео систематску анализу географске пристрасности NER модела при препознавању адресних израза. Формиран је одговарајући скуп података, реализована је контролисана експериментална поставка и извршена анализа резултата на нивоу држава, градова и компоненти сложених адресних структура. Добијени резултати показују да географска пристрасност представља сложену појаву условљену географским пореклом података, врстом адресне компоненте, структурним контекстом и конкретним моделом.

Кандидат је исказао самосталност и систематичност у истраживачком поступку, као и способност да формулише релевантне истраживачке хипотезе, реализује експерименте над обимним скупом података и критички протумачи добијене резултате. Рад представља допринос анализи пристрасности NLP модела за српски језик и отвара простор за даља истраживања у овој области.

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

Београд, 12.09.2026. године

Чланови комисије:

др Бошко Николић Редовни професор
сагласан, 11.09.2026.

др Вук Батановић Виши научни сарадник
сагласан, 12.09.2026.

др Владимир Јоцовић Доцент
сагласан, 11.09.2026.