

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. Михаило Жарковић под насловом „Анализа пристрасности модела за препознавање именованих ентитета у српском језику према националним мањинама". Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Средње образовање завршио сам у Петој београдској гимназији, на природно-математичком смеру (2014–2018). Након тога, уписао сам основне студије Софтверског инжењерства на Електротехничком факултету Универзитета у Београду, које сам успешно завршио 2022. године.

Тренутно сам студент мастер студија на истом смеру, где се додатно усавршавам у областима напредног софтверског инжењерства и технолошких иновација.

Поред академског образовања, већ три године радим као софтверски инжењер.

2. Извештај о студијском истраживачком раду

Кандидат Михаило Жарковић је као припрему за израду мастер рада урадио истраживање релевантне литературе која се односи на препознавање именованих ентитета. Посебна пажња је дата анализи пристрасности код препознавања имена националних мањина. Конкретно, анализирани су постојећи модели за српски језик и различити приступи у анализи пристрасности. Одабрано је шест модела за српски језик: ICEF-NLP/bcms-bertic-comtext-sr-legal-ner-ekavica, ICEF-NLP/bcms-bertic-comtext-sr-legal-ner-ijekavica, classla pipeline за стандардни српски језик, classla pipeline за нестандартни српски језик и SrpCNER. Модели се разликују по архитектурама и по корпусима на којим су тренирани.

3. Опис мастер рада

Мастер рад обухвата 46 страна, са укупно 4 слике, 30 табела и 37 референци. Рад садржи увод, 5 поглавља и закључак и списак коришћене литературе.

Прво поглавље представља увод у коме су описани предмет и циљ рада. Представљена је савремена архитектура модела и значај овакве врсте анализе у процесу интеграције свих националних мањина. Друго поглавље садржи опис коришћених модела и преглед различитих врста анализе пристрасности. Дат је опис модела, њихове карактеристике и увид у корпусе на којима су тренирани. Посебна пажња је дата приступима у анализи пристрасности из којих је формирана методологија коришћена у раду.

Треће поглавље описује методологију рада. Представљена су два комплементарна режима евалуације: без и са контекстом. Објашњено је шта контекст представља, на који начин се формира и које се метрике користе. Четврто поглавље приказује процес формирања корпуса имена, како српских тако и имена националних мањина. Акценат је стављен и на формирање скупа одабраних реченица за потребе обогаћивања контекста модела. Детаљно су описани критеријуми и ограничења примењени у процесу сакупљања података.

Пето поглавље описује технологију и програмску имплементацију експеримената. Анализиране су перформансе модела у временском домену, зависност од хардвера на којем се покрећу и њихове саме архитектуре.

У оквиру шестог поглавља представљени су добијени резултати подељени по типу експеримента. Унутар експеримената резултати су анализирани на нивоу модела, националне мањине, домена контекста и начина навођења личног имена. Детекције српских имена у оба експеримента су коришћене као базични резултати. Детаљно су анализирани резултати модела који уносе пристрасност у своје детекције.

Седмо поглавље садржи закључак рада. У њему су сумиране најчешће грешке модела, наведен је модел са најлошијим перформансама, продискутовани су могући разлози његовог понашања и предложене смернице за наставак рада у овој области.

4. Анализа са кључним резултатима

Мастер рад Михаила Жарковића се бави анализом пристрасности модела за препознавање именованих

ентитета у српском језику према националним мањинама. Примењена методологија и формирано скупови података могу бити употребљени и за анализе других врста пристрасности.

У првом експерименту, без контекста, је утврђено да модели често имају проблем да препознају ентитете када се на улазу нађе само име, а нарочито само презиме. Модели посебно показују пристрасност према одређеним националним мањинама попут: албанске, мађарске, ромске и словачке.

У другом експерименту, са контекстом, удео пристрасности се доста смањивао али се може рећи да у одређеним доменима препознавање презимена и даље представља проблем за поједине националне мањине. Модел SrpCNER се у оба експеримента издвојио као најбољи, разлози за то се могу тражити како у архитектури модела тако и у корпусу за тренирање.

Основни доприноси рада су: 1) формирано корпуси имена и презимена националних мањина; 2) дефинисан систематичан приступ за анализу пристрасности NER модела; 3) могућност поновне употребе корпуса и методологије у испитивањима других пристрасности.

5. Закључак и предлог

Кандидат Михаило Жарковић је у свом мастер раду успешно реализовао анализу пристрасности модела за препознавање именованих ентитета у српском језику према националним мањинама. Кандидат је представио добијене резултате, изнео њихово тумачење и дао предлоге за наставак истраживања у овој области, а све то у циљу што боље интеграције свих припадника националних мањина.

Кандидат је исказао самосталност и систематичност у своме поступку као и иновативне елементе у решавању проблематике овог рада.

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

Београд, 26.10.2025. године

Чланови комисије:

др Бошко Николић Редовни професор
сагласан, 26.10.2025.

др Вук Батановић Научни сарадник
сагласан, 26.10.2025.

др Владимир Јоцовић Доцент
сагласан, 26.10.2025.