

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 02.03.2021. године именовало нас је у Комисију за преглед и оцену мастер рада дипл. инж. Иване Муњас под насловом „Класификација адреса помоћу обраде текста и машинског учења”. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Ивана Муњас је рођена 29.08.1996. године у Смедереву. Завршила је основну школу “Др Јован Цвијић” у Смедереву као вуковац. Уписала је Гимназију у Смедереву коју је завршила са одличним успехом. Електротехнички факултет уписала је 2015. године. Дипломирала је на одсеку за Сигнале и системе 2019. године са просечном оценом 9,24. Дипломски рад одбранила је у септембру 2019. године са оценом 10. Током основних студија учествовала је на бројним радионицама и такмичењима. Радила је као студент демонстратор, а ван школских активности похађала је разне летње истраживачке програме. Дипломске академске – мастер студије на Електротехничком факултету у Београду, на Модулу за сигнале и системе уписала је у октобру 2019. године. Положила је све испите са просечном оценом 10.

2. Извештај о студијском истраживачком раду

Кандидат Ивана Муњас је као припрему за израду мастер рада урадила истраживање релевантне литературе која се односи на област којој припада тема мастер рада. Конкретно, анализирана су постојећа решења и проблеми у области класификације адреса и кратког текста. Истраживањем области утврђено је да се конкретан проблем класификације не појављује у литератури, али да се у сличним радовима користи: претпроцесирање текста, векторизација текста углавном помоћу вектора значења речи и различити модели машинског учења за класификацију. Анализом решења је утврђено да треба испратити све поменуте кораке, али је у овом раду за сваки корак испробано више различитих метода да би се дошло до најбољег могућег решења.

3. Опис мастер рада

Мастер рад обухвата 44 стране са укупно 18 слика, 20 табела и 8 референци. Рад садржи увод, 3 поглавља и закључак (укупно 5 поглавља) и списак коришћене литературе.

Прво поглавље представља увод у коме су описани предмет и циљ рада. Представљени су најчешћи проблеми у обради природних језика и објашњено је какве конвенције важе за америчке адресе и која ће правила бити усвојена за припадност адреса одговарајућим класама.

У другом поглављу је дат кратак преглед радова који се односе на класификацију адреса, упаривање адреса и класификацију кратког текста.

У трећем поглављу је детаљно описан метод рада. Описано је како су подаци прикупљени и лабелирани, затим на који начин се вршило претпроцесирање адреса. Описана су два начина векторизације текста: приступ вреће речи и вектори значења речи. На крају су описани и класификациони модели који су коришћени: логистичка регресија, метода потпорних вектора, метода случајне шуме, наивни Бајесов класификатор и неурална мрежа.

Четврто поглавље представља резултате на валидационом и тест скупу података. Рад свих поменутих модела прво је анализиран у комбинацији са приступом вреће речи, а затим са векторима значења речи. Модели који су се најбоље показали на валидационом скупу су тестирани на тест скупу како би се добила финална процена рада модела.

Шесто поглавље представља закључак и дискусију у оквиру кога је описан значај описаног решења и могућа даља унапређења. Резимирани су резултати рада и анализирани су доступни резултати из литературе.

4. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Иване Муњас се бави проблематиком класификације америчких адреса у шест различитих класа. Класификациони модели који се баве адресама налазе широку примену у бројним интернет сервисима, као што су сервиси за продају добара или одређивање локације.

Оваквом класификацијом се омогућава да се адресе које не испуњавају одређене стандарде искључе или исправе на одговарајући начин. С обзиром да су адресе анализирани у овом раду добијене од веб-сајта за продају некретнина, њихова тачност омогућава корисницима да лакше пронађу жељену некретнину.

Основни доприноси рада су: 1) приказ и методологија креирања класификационих модела и метода претпроцесирања текста; 2) примена креираних модела у оквиру различитих система који користе адресе и за које је важна њихова тачност; 3) могућност наставка рада ради креирања још квалитетније класификације.

5. Закључак и предлог

Кандидат Ивана Муњас је у свом мастер раду успешно решила проблем класификације америчких адреса и креирала је низ метода које се користе за претпроцесирање и векторизацију текста. Приказала је упоредно поређење различитих метода векторизације текста и коришћења различитих модела машинског учења. Предложена побољшања могу значајно да унапреде квалитет класификације и да дају додатну информацију за упоредну анализу различитих модела.

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

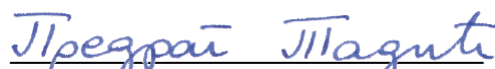
На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Иване Муњас прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 02.09.2021. године

Чланови комисије:



др Бошко Николић, ред. професор



др Предраг Тадић, доцент