

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Одговарајуће комисије Електротехничког факултета у Београду и Факултета организационих наука у Београду именовале нас су у Комисију за преглед и оцену мастер рада мастер инж. Марије Костић под насловом „Једнојезична, вишејезична и међујезична класификација коментара из програмског кода“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидаткиње

Марија Костић је рођена 23.05.1997. године у Београду. Основну школу „Аца Милосављевић“ у Београду завршила је као носилац Вукове дипломе и Ђак генерације. Након тога је уписала Прву спортску кошаркашку гимназију у Београду коју је завршила такође као Ђак генерације. Електротехнички факултет у Београду уписала је 2015. године. Дипломирала је као најбољи студент на Одсеку за софтверско инжењерство 2019. године са просечном оценом 9,96. Дипломски рад на тему „Софтверски систем за визуелизацију података коришћењем мапа“ одбранила је у септембру 2019. године са оценом 10. Додељена јој је награда „Гордана Јокић Кашиковић и Драгиша Кашиковић“. Мастер академске студије на Електротехничком факултету у Београду је уписала октобра 2019. године на Модулу за софтверско инжењерство. Положила је све испите са просечном оценом 10,0 и одбранила је мастер рад на тему „Примена техника машинског учења и анализе социјалних мрежа у предвиђању исхода тениских мечева“ у септембру 2020. године са оценом 10. Докторске академске студије на Електротехничком факултету у Београду је уписала 2020. године. Положила је све испите са оценом 10.

Током студија радила је као студент демонстратор при Катедри за рачунарску технику и информатику. Такође је обавила и више стручних пракси као софтверски инжењер, и то: 2019. године у Развојном центру компаније Мајкрософт у Београду и у компанији Гугл у Цириху, 2021. године у компанији Гугл у Амстердаму и 2022. године у компанији Квадратур у Лондону и компанији Гугл у Цириху. Два пута је представљала факултет на Међународним сусретима студената електротехнике – Електријади. Од марта 2021. је запослена у Иновационом центру Електротехничког факултета у Београду као истраживач. Објавила је четири научна рада на међународним конференцијама.

Мастер академске студије у оквиру Мастер 4.0 програма Напредне информационе технологије у дигиталној трансформацији на Модулу анализа података уписала је у октобру 2019. године. Положила је све испите са просечном оценом 9,9.

2. Извештај о студијском истраживачком раду

Кандидаткиња Марија Костић је као припрему за израду мастер рада урадила истраживање релевантне литературе која се односи на област којој припада тема мастер рада. Конкретно, анализирала је постојећа решења и проблеме у области класификације коментара из програмског кода. Истраживањем области утврдила је да се постојећи системи за класификацију коментара баве само коментарима из најпопуларнијих програмских језика (Java, C/C++, Python) писаних само на енглеском језику. Сви ти системи су базирани на традиционалним моделима машинског учења. Даљим истраживањем литературе у области дубоког учења је утврдила да модели машинског учења засновани на Трансформер архитектури представљају перспективно решење. Неки од њих су и дизајнирани да подрже

рад са подацима на више различитих природних језика (енглеском и српском за потребе овог истраживања).

3. Опис мастер рада

Мастер рад обухвата 36 страна са укупно 14 слика, 5 табела и 34 референце. Рад садржи увод, 4 поглавља и закључак (укупно шест поглавља), списак коришћене литературе, затим списак скраћеница, списак слика и списак табела.

Прво поглавље представља увод са анализом мотивације за рад на описаном предмету истраживања. Описани су предмет и циљ рада, као и његови главни доприноси.

У другом поглављу је дат преглед литературе која се бави класификацијом коментара из програмског кода. За сваки рад су дати кратки описи коришћених скупова података и наведени коришћени класификациони модели заједно са постигнутим резултатима.

Треће поглавље описује процес припреме података који су касније коришћени у истраживању и експериментима. Објашњен је начин прикупљања коментара и њихове анотације. Поред тога је приказана и основна анализа коначног скупа података.

У четвртом поглављу је детаљно изложена методологија рада. Након прецизне дефиниције проблема који се решава, дате су теоријске основе за моделе машинског учења који су коришћени – Метод потпорних вектора и модели засновани на Трансформер архитектури. На крају је описан процес евалуације.

Пето поглавље садржи детаље о резултатима свих спроведених експеримената. Прво су приказани резултати једнојезичне класификације, затим вишејезичне класификације и на крају међујезичне класификације. Сви ови резултати су на крају упоређени и продискутовани.

Последње, шесто поглавље је закључак у оквиру кога је урађена рекапитулација истраживања. Резимирани су резултати рада, наведена нека његова ограничења и дате смернице за будући рад.

4. Анализа рада са кључним резултатима

Мастер рад мастер. инж. Марије Костић се бави проблематиком класификације коментара из програмског кода. Разматрани коментари су извучени из кода који је писан на чак осам различитих програмских језика и два природна језика. Реализовани класификатори даље могу да се користе као полазни корак у решавању комплекснијих проблема у софтверском инжењерству као што је семантичка претрага програмског кода.

Главни доприноси рада су:

- Детаљан преглед научних радова у области класификације коментара из програмског кода.
- Прво истраживање које се бави вишејезичном и међујезичном класификацијом коментара.
- Нова таксономија за класификацију коментара која може да се примени на различите програмске језике.
- Нови јавно доступни скуп од око 10.000 коментара преузетих из осам програмских језика (C, C++, C#, Java, JavaScript/TypeScript, PHP, Python и SQL). Половина коментара је писана на српском језику, док је друга половина писана на енглеском језику.
- Прво истраживање које се бави аутоматизованом класификацијом коментара на српском језику.

5. Закључак и предлог

Кандидаткиња Марија Костић је у свом мастер раду успешно решила проблем класификације коментара у програмском коду. Урађено истраживање је прво истраживање које се бави класификацијом коментара из различитих програмских и природних језика. Предложене смернице за даљи рад могу да унапреде перформансе модела. Развијени класификатори се свакако могу користити као први корак у решавању комплекснијих проблема у софтверском инжењерству.

При реализацији истраживања, кандидаткиња Марија Костић је показала значајан степен аналитичности, систематичности и одговорила је на све захтеве који су јој били постављени. Исказала је самосталност у раду и предложила је иновативна елементе за решавање проблематике овог рада.

На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад мастер инж. Марије Костић прихвати као мастер рад и кандидаткињи одобри јавну усмену одбрану.

Београд, 18.04.2023. године

Чланови комисије:

Б. Николић

Др Бошко Николић, ред. проф.

Б. Делибашић

Др Борис Делибашић, ред. проф.

П. Тадић

Др Предраг Тадић, доцент