

## ИЗВЕШТАЈ

о прегледу мастер рада

### „Класификација текста према сентименту применом метода машинског учења”

кандидат: Марија Радовић

Одлуком Наставно-научног већа Математичког факултета која је донета на 370. редовној седници одржаној 24.04.2020. (електронска седница) године именовани смо за чланове комисије за преглед и оцену мастер рада под насловом „Класификација текста према сентименту применом метода машинског учења” кандидаткиње Марије Радовић, број инекса 1065/2018, студента мастер студија на студијском програму Математика, модул Рачунарство и информатика на Математичком факултету.

## 1 Област рукописа

Рукопис „Класификација текста према сентименту применом метода машинског учења” кандидаткиње Марије Радовић, припада следећим областима: обрада природних језика, машинско учење.

## 2 Структура рукописа и кратак приказ

Рукопис се састоји од 47 страна, организованих у 7 поглавља и библиографију.

У оквиру главе „Увод” дат је кратак увод у област којом се рад бави и наведени су циљеви и мотивација рада.

Друга глава рада „Решавање проблема методама машинског учења” описује процес решавања проблема сентимент анализе применом метода машинског учења. Састоји се од поглавља 2.1 Претпроцесирање текста, 2.2 Креирање векторског модела, 2.3 Тренирање модела, 2.4 Евалуација модела и 2.5 Побољшање модела.

У оквиру треће главе „Векторске репрезентације текста” даје се детаљан приказ различитих начина репрезентације текста који су коришћени у овом раду. Сваком од начина посвећено је по једно поглавље: 3.1 Врећа речи, 3.2 *N*-грамски модели, 3.3 Вектор врста речи, 3.4 Вектор позиција речи у тексту, 3.5 *TF* модел, 3.6 *TF-IDF* модел и 3.7 *Word2Vec* модел.

Четврта глава „Методе машинског учења” даје опис коришћених метода машинског учења. Састоји се из 5 поглавља: 4.1 Метод потпорних вектора, 4.2 Наивни Бајесов класификатор, 4.3 Наивни Бајесов класификатор са методом потпорних вектора, 4.4 Потпуно повезане неуронске мреже

У оквиру пете главе „Имплементација” дати су детаљи имплементације целог система.

Шеста глава „Резултати” описује јавно доступне корпусе филмских рецензија на српском, турском и енглеском језику, као и резултате добијене применом развијених метода на овим корпусима. Састоји се из 4 поглавља: 6.1 Корпус на српском језику, 6.2 Корпус на енглеском језику, 6.3 Корпус на турском језику и 6.4 Емпиријско подешавање и резултати.

Основни закључци овог рада изнети су у глави „Закључак”.

Глава „Библиографија” садржи списак од 18 библиографских јединица које је кандидаткиња користила приликом писања рада.

### 3 Анализа рукописа

У рукопису који смо анализирали, кандидат даје адекватан увод у основне појмове који се користе у раду. Описује проблем класификације текста према сентименту, различите моделе репрезентације текста, као и различите методе машинског учења коришћене у циљу класификације текста. Основни допринос овог рада је одговарајућа имплементација система за класификацију текста према сентименту применом различитих метода машинског учења на корпусе филмских рецензија на српском, енглеском и турском језику. Овај систем је јавно доступан на адреси <https://github.com/mradovic23/ML-SPD> и може се користити као добра основа за даља истраживања. Резултати овог рада су приказани на међународној INISTA 2020 конференцији (<http://inista.org/>) и биће штампани у зборнику ове конференције.

### 4 Закључак и предлог

Реализацијом овог рада и пратећом имплементацијом, кандидаткиња Марија Радовић је показала да је у стању да самостално овлада знањима из једне научне дисциплине, имплементира потребне методе и спроведе одговарајућу експерименталну евалуацију имплементираних метода. Тиме је кандидаткиња показала да у потпуности задовољава захтеве који се постављају у изради мастер рада. На основу свега наведеног Комисија предлаже да се рукопис под насловом:

**„Класификација текста према сентименту применом метода машинског учења“**

кандидаткиње Марије Радовић прихвати као мастер рад и да се закаже његова јавна усмена одбрана.

Београд, 31.08.2020.

Комисија:

доц. др Јелена Граовац, ментор

доц. др Младен Николић

доц. др Јована Ковачевић