

ИЗВЕШТАЈ

о прегледу мастер рада

„Примена лексичких ресурса у сентимент анализи текста”

кандидат: Јелена Ћосић

Одлуком Наставно-научног већа Математичког факултета која је донета на 368. редовној седници одржаној 24.01.2020. године именовани смо за чланове комисије за преглед и оцену мастер рада под насловом „Примена лексичких ресурса у сентимент анализи текста” кандидаткиње Јелене Ћосић, број инекса 1081/2018 студента мастер студија на студијском програму Математика, модул Рачунарство и информатика на Математичком факултету.

1 Област рукописа

Рукопис „Примена лексичких ресурса у сентимент анализи текста” кандидаткиње Јелене Ћосић, припада области обраде природних језика.

2 Структура рукописа и кратак приказ

Рукопис се састоји од 43 стране, организоване у 6 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 „Лексички ресурси” описује лексичке ресурсе који су у овом раду коришћени за решавање проблема класификације текста према сентименту. Састоји се од поглавља 2.1 *Wordnet* и 2.2 *Електронски речник за српски језик*.

У оквиру треће главе „Сентимент класификација текста” даје се приказ проблема класификације текста према сентименту, као и опис неких метода које се могу користити у циљу решавања проблема класификације текста. Приказане су и методе које ће бити коришћене за евалуацију добијених резултата класификације. Састоји се из следећих поглавља: 3.1 *Препроцесирање текста*, 3.2 *Методе класификације текста*, 3.3 *Евалуација квалитета класификације*.

Четврта глава „Развијене методе сентимент класификације” даје детаљан опис метода развијених у циљу решавања проблема класификације текста према сентименту. Описује корпуре филмских рецензија на српском и енглеском језику који су коришћени у овом раду, као и детаље имплементације самог система. Састоји се из 5 поглавља: 4.1 *Методе засноване на лексичким ресурсима*, 4.2 *Хибридне методе*, 4.3 *Корпуси филмских рецензија*, 4.4 *Python и корисни пакети* и 4.5 *Имплементација*.

У оквиру пете главе „Резултати” приказани су резултати добијени применом описаних метода на корпуре на српском и енглеском језику. Сваком од корпуса посвећено је по једно поглавље: 5.1. Резултати за корпус на енглеском језику и 5.2. Резултати за корпус на српском језику.

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

Глава „**Библиографија**” садржи списак од 21 библиографских јединица које је кандидаткиња користила приликом писања рада.

3 Анализа рукописа

У рукопису који смо анализирали, кандидаткиња даје увод у основне појмове који се користе у раду. Описује проблем сентимент анализе текста и нека од могућих решења овог проблема применом различитих метода заснованих на лексичким ресурсима. Основни допринос овог рада је одговарајућа имплементација система за класификацију текста према сентименту применом различитих метода на јавно доступне корпусе филмских рецензија на српском и енглеском језику. Овај систем је јавно доступан на адреси <https://github.com/JelenaCosic1994/sentiment-text-analysis-using-lexical-resources> и може се користити као добра основа за даља истраживања.

4 Закључак и предлог

Реализацијом овог рада и пратећом имплементацијом, кандидаткиња Јелена Ћосић 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 задовољава захтеве који се постављају у изради мастер рада. На основу свега наведеног Комисија предлаже да се рукопис под насловом:

„Примена лексичких ресурса у сентимент анализи текста“

кандидаткиње Јелене Ћосић прихвати као мастер рад и да се закаже његова јавна усмена одбрана.

Београд, 4.09.2020.

Комисија:

доц. др Јелена Граовац, ментор

проф. др Гордана Павловић-Лажетић

проф. др Цветана Крстев