

**МОЛБА
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА**

Молим да ми се одобри израда мастер рада под насловом:

„Имплементација алата за препроцесирање и одређивање растојања података”

Значај теме и области:

Препроцесирање података је један од важних корака у процесу истраживања података. Применом различитих метода препроцесирања података обезбеђују се неопходни предуслови за коректну примену алгоритама истраживања података (за класификацију, кластеровање, правила придруживања). Сви алати за истраживање података и машинско учење имају имплементирани неке од метода препроцесирања података. Међутим, ниједан од алата нема имплементиран комплетан скуп процедура за препроцесирање, или је примена таквих процедура ограничена карактеристикама података (нпр. димензионалност, кардиналност или врста типа улазних података). Такође, сваки алат садржи неки скуп мера (метрика) које су придружене неком од подржаних алгоритама. Међутим, као и у претходном случају, број мера/метрика који је подржан је релативно мали што у неким случајевима онемогућава добијање квалитетних резултата. Недостатак одређених процедура за препроцесирање и расположивост метрика представља ограничавајући фактор посебно у случају када је потребно применити неки од нестандартних поступака или алгоритама који су потребни у истраживању али нису имплементирани у комерцијалним алатима.

Специфични циљ рада:

Циљ рада је имплементација широког скупа метода препроцесирања податка и метрика растојања који поред уобичајеног скупа техника (транспонованање, дискретизација, бинаризација, нормализација, рад са непознатим вредностима, итд.) и метрика (растојање Минковског, Жакардово растојање, Махаланобисово растојање, Ф-мера, косинусно растојање, итд.) укључује и методе које нису имплементирани у оквиру алата расположивих на тржишту (издвајање/искључивање подскупова и поједине групе података, издвајање/искључивање атрибута са специфичном вредношћу, итд.). Такође, резултат рада ће бити и имплементација скупа мера/метрика за различите типове података које нису уобичајено имплементирани у расположивим алатима као и Lance-Williams-ове формула за сличност кластера.

Ненад Перишић, 1105/2019, Информатика

(име и презиме студента, бр. индекса, модул)

Сагласан ментор

Ненад Митић

(својеручни потпис студента)

(својеручни потпис ментора)

29.04.2021.

(datum подношења молбе)

Чланови комисије

1. Саша Малков
2. Александар Картељ

Катедра **за Рачунарство и информатику**

је сагласна са предложеном темом.

(шеф катедре)

(datum одобравања молбе)