

МОЛБА
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА

Молим да се одобри израда мастер рада под насловом:

„Примена лексичких ресурса у сентимент анализи текста”

Значај теме и области:

Сентимент анализа је сложен аналитички процес који се реализује кроз примену различитих метода обраде природног језика, машинског учења и рачунарске лингвистике у циљу систематског идентификовања, добијања, квантификовања и анализе емоција и мишљења исказаних у тексту. Примери њене широке примене су анализа мишљења о разним производима и услугама добијених на различитим анкетама, друштвеним мрежама, форумима и блогovima, са циљем унапређења доношења одлука у трговини, финансијском сектору, политици и у животу уопште. Један од задатака сентимент анализе је класификација докумената према поларитету исказаног мишљења у тексту (позитивно, негативно или неутрално). Решавању проблема класификације може се приступити применом метода заснованих на лексичким ресурсима, применом метода машинског учења или хибридном приступом.

Специфични циљ рада:

Циљ овог рада је сентимент класификација јавно доступних корпуса филмских рецензија на српском и енглеском језику, коришћењем метода заснованих на лексичким ресурсима. Користиће се два приступа: приступ заснован на методи ненадгледаног и приступ заснован на методи надгледаног учења. Оба приступа ће се заснивати на примени српског и енглеског SentiWordnet-а, лексичко-семантичке мреже за српски и енглески језик. Проблем флексије у српском језику биће решен коришћењем морфолошког речника за српски језик. Биће извршена анализа и интерпретација добијених резултата, њихово поређење и испитивање да ли се укључивањем морфолошких, синтаксичких и семантичких информација садржаних у лексичким ресурсима може унапредити процес сентимент анализе на српском (као једном од морфолошки богатијих језика) и енглеском (као најраспрострањенијем језику на свету).

Остале битне информације:

Апликација ће бити развијана у програмском језику *Python* и биће јавно доступна.

Литература:

1. Manning, Christopher D., and Hinrich Schtze., *Foundations of statistical natural language processing*, Vol. 999. Cambridge: MIT press, 1999.
2. Denecke, Kerstin, *Using SentiWordnet for multilingual sentiment analysis*. *Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on*, IEEE, 2008.

Јелена Ђосић, 1081/2018, МР

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

Сагласан ментор др Јелена Граовац, доцент

_____ (својеручни потпис студента)

_____ (својеручни потпис ментора)

_____ (датум подношења молбе)

Чланови комисије

1. проф. др Цветана Крстев
редовни професор Филолошког факултета
2. проф. др Гордана Павловић-Лажетић

Катедра _____ је сагласна са предложеном темом.

_____ (шеф катедре)

_____ (датум одобравања молбе)