

Катедри за рачунарство и информатику

Предмет: Сагласност за одбрану мастер рада

Одлуком Катедре и ННВ од 09.05.2014. године именовани смо у комисију за одбрану мастер рада под насловом *"Условна случајна поља и примена на растављање речи на крају реда у српском језику"* кандидата Иване Медић, студента мастер студија на студијском програму *Математика-Рачунарство и информатика* на Математичком факултету. Тема рада је метода машинског учења структурне предикције - *условних случајних поља* - заснована на вероватносном графовском моделу и њена примена у решавању проблема хифенације – растављања речи на крају реда у српском језику.

Рад се састоји од четири поглавља и закључка. Први део је Увод и у њему се уводе основни појмови везани за методу машинског учења која се излаже у раду, проблем хифенације у српском језику као и начин на који ће се овај проблем решавати у раду применим изложене методе. Приказује се у кратким цртама и садржај рада.

У поглављу 2 приказује се метода условних случајних поља, посебно њена најједноставнија али и најчешће коришћена варијанта – линеарна условна случајна поља, као и уопштење – генерална условна случајна поља. Описани су и пробабилистички модели који се могу посматрати као основа за развој условних случајних поља - дискриминативни модели линеарне и логистичке регресије, модел максималне ентропије, скривени Марковљеви модели и Марковљеви модели максималне ентропије, као и однос који међу овим моделима постоји. У овом поглављу изложени су и алгоритми за обучавање параметара предикције, као и алгоритми за решавање проблема закључивања неопходни за учење параметара.

Поглавље 3 описује проблем растављања речи на крају реда са којим се срећемо при прелому текста и припреми за приказ или штампу. Приказан је кратки осврт на развој метода и софтвера за поделу речи на слоге у разним језицима, као и правила за растављање речи на крају реда у српском језику.

У поглављу 4 представљена је примена условних случајних поља на решавање проблема растављања речи на крају реда у српском језику. Представљен је начин кодирања секвенци (речи), описане су карактеристичне функције које представљају основу модела као и параметри модела и скупови за обучавање и тестирање. За имплементацију модела искоришћен је софтвер CRF++, елементи модела су трансформисани у облик који овај софтвер захтева, извршено је обучавање параметара и сама предикција. На крају поглавља анализирани су добијени резултати. Прецизност модела је висока и на скупу од 31000 речи растављених постојећим софтвером и подељених у односу 80:20 на скупове за обучавање и тестирање, прелази 99% на нивоу слова.

У Закључку се сумирају главни елементи рада. Следи списак од 30 коришћених референци.

Закључак Комисије

Увидом у текст *„Условна случајна поља и примена на растављање речи на крају реда у српском језику“* мишљења смо да је кандидат, Ивана Медић, овладала сложеном облашћу метода машинског учења у обиму који омогућује креативну и нетривијалну примену на решавање реалних отворених проблема. Појмови и алгоритми везани за изложену методу и моделе су изложени егзактно и прецизно и сведоче о оствареној темељној и суштинској надградњи програма мастер студија рачунарства. Резултати добијени примењеном методом обећавају нови истраживачки домет. Са задовољством предлагемо Катедри да одобри јавну одбрану овог мастер рада.

У Београду,
16. септембра 2014.

Комисија

1. _____
/др Гордана Павловић-Лажетић, р.проф./

2. _____
/др Душко Витас, в. проф./

3. _____
/др Младен Николић, доцент/