

МОЛБА
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА

Молим да се одобри израда мастер рада под насловом:

„Аутоматска транскрипција говора за српски језик”

Значај теме и области:

Аутоматска транскрипција говора данас представља једну од битних области обраде природног језика, са директним утицајем на дигиталну приступачност и ефикасност управљања информацијама. У ери у којој се генеришу огромне количине аудио и видео садржаја, способност прецизног претварања говора у текст постаје неопходна за претрагу мултимедијалних база података, титловање у реалном времену и подршку особама са оштећењем слуха. Развој робусних модела који су отпорни на буку, различите акценте и специфичне термине од кључног је значаја за даљу интеграцију вештачке интелигенције у свакодневне радне процесе, од правосуђа до медицине.

Специфични циљ рада:

Примарни циљ овог истраживања је компаративна евалуација перформанси савремених архитектура за препознавање говора, са фокусом на различите верзије OpenAI/Whisper модела. Рад тежи да идентификује оптималне конфигурације хиперпараметара, као што су брзина учења и регулација тежина, кроз процес финог подешавања на специфичним скуповима података. Поред саме транскрипције, посебан циљ је испитивање ефикасности пост-процесирања текста путем хибридног алгорита за проверу правописа, како би се квантификовало у којој мери статистичка и лингвистичка корекција могу компензовати грешке настале у фази акустичког моделовања.

Остале битне информације:

За техничку реализацију предложеног истраживања биће коришћен програмски језик Python заједно са библиотеком PyTorch, која омогућава хардверски убрзану обуку и извршавање дубоких неуронских мрежа. Поступак претпроцесирања транскрипција обухватиће примену библиотеке num2words за нормализацију речи које представљају нумеричке вредности, као и библиотеке transliterate за пребацивање текста у српску ћирилицу. Такође, биће примењене методе паралелизације процеса ради ефикасне обраде обимних скупова података. Евалуација модела вршиће се кроз стандардизоване метрике WER и CER.

Стефан Јевтић, 1043/2024, Информатика

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

Сагласан ментор проф. др Младен Николић

(својеручни потпис студента)

26.01.2026.

(datum подношења молбе)

(својеручни потпис ментора)

Чланови комисије

- проф. др Јелена Граовац
- проф. др Јована Ковачевић

Катедра за рачунарство и информатику

је сагласна са предложеном темом.

(шеф катедре)

(datum одобравања молбе)