

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 15.03.2022. године именovalo нас је у Комисију за преглед и оцену мастер рада дипл. инж. Драга Албијанића под насловом „Употреба конвулативне неуралне мреже у сврхе процене музичког жанра“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Драго Албијанић је рођен 26.10.1994. године у Београду. Завршио је основну школу „Јосип Панчић“ у Београду као носилац Вукове дипломе. Уписао је Тринаесту београдску гимназију у Београду коју је завршио са одличним успехом. Електротехнички факултет је уписао школске 2013/2014. године. Дипломирао је на одсеку за Софтверско инжињерство у септембру 2020. године са просечном оценом 7.16. Дипломски рад је одбранио у септембру 2019. године са оценом 10. Тема дипломског рада је била „Интернет апликација за управљање пословима“, под менторством проф. др Милоша Цветановића. Исте године је на истом факултету уписао дипломске академске -мастер студије (Модул софтверско инжињерство) на којима је последњи испит положио јуна 2021. године. Положио је све испите са просечном оценом 10.

2. Извештај о студијском истраживачком раду

Кандидат Драго Албијанић је као припрему за израду мастер рада „Употреба конвулативне неуралне мреже у сврхе процене музичког жанра“ урадио истраживање релевантне литературе која се односи на проблем практичне употребе конвулативних неуралних мрежа. Кандидат је разматрао постојећа решења у домену процене музичког жанра. Такође, кандидат је обавио претрагу доступних скупова података који се користе за тестирање перформанси постојећих решења.

3. Опис мастер рада

Мастер рад има 82 стране основног текста и 45 страна прилога, односно 127 страна укупно. Рад садржи 104 слике и 88 референци. Рад садржи увод, 8 поглавља и закључак (укупно 9 поглавља), списак коришћене литературе и списак слика.

Прво поглавље представља увод у коме су описани предмет, садржај и циљ рада, као и историјски контекст у коме се развијала проблематика која се налази у средишту рада.

Друго поглавље даје кратак опис карактеристика и природе звука као таласног кретања као и опис представљања звука у временском и фреквенцијском домену. Такође је описан начин слагања звучних таласа са описом настанка појаве интерференције, модулације и сличних ефеката у преносу и пријему звучних таласа.

Треће поглавље се односи на физиолошку акустику. У најкраћем је описан рад органа чула говора и чула слуха код човека са приказом њихових виталних делова. Дате су основне карактеристике звука изражене кроз његову јачину, висину и боју.

Четврто поглавље се односи на математичку интерпретацију Фуријеових трансформација (*FT*) на којима је базирана анализа звука и музике кроз декомпозицију сложених сигнала на основне компоненте од којих је аудио сигнал састављен. Такође је описан и концепт Краткотрајне Фуријеове трансформације (*STFT*). На крају је дат приказ Спектограма који представља полазну основу у процесу који се односи на аутоматско препознавање музике (*Automatic Music Genre Recognition - AMGR*).

Пето поглавље систематски приказује детаљан поступак креирања *MFCC* кепстралних коефицијената који представљају основно звучно обележје делова песме, односно целе песме. Комплетан процес се изводи на бази коришћења *Python* библиотеке *librosa*.

Шесто поглавље описује структуру и начин функционисања конволутивне неуралне мреже. Основни циљ овог поглавља је да се читаоцу појасни трансформација од матричног улаза, у виду слике, до саме предикције.

Седмо поглавље тј. први део овог поглавља описује настанак новог скупа података, који је састављен првенствено за потребе овог рада. У њему је наведен део кода који је задужен за скидање жанровских листи са *YouTube* сајта. Новонастали скуп обухвата 13 музичких жанрива. Други део овог поглавља представља својеврстан историјат сваког од ових жанрова.

Осмо поглавље се бави практичним делом овог рада, тј. имплементацијом и резултатима два модела који су тренирани на *GTZEN* скупу и новонасталом скупу података. Приказани су графици прецизности и губитка, матрице конфузије, извештај класификације као и тест који пореди два модела.

Девето поглавље је упутство за употребу интернет апликације која служи да се визуализују резултати предикција ових модела. Сервер стана је реализована у *FastAPI* технологији док се на клијентској страни користи *Chart.js* библиотека.

Да би главни део рада био растерећен детаљима у структуру рада су уведени прилози *III-IV* који су дати на његовом крају у којима је садржано много више детаља из одређених области.

Прилог П1 се односи на приказ музичких инструмената и приказ основних карактеристика музике. У Прилогу П2 је описан појам, значај и примена конволуцијског оператора. У Прилогу П3 су дате основе временске дискретизације сигнала и шире дигиталне обраде сигнала у временском и фреквенцијском домену. Прилог П4 је резултат кандидатовог рада у фирми која се професионално бави испитивањима динамичких феномена сложених техничких система која у својој основи између осталог има снимање и анализу сигнала вибрација и звука са комплетном мерно-рачунарском опремом која се користи у ту сврху. У Прилогу П5 су дате листе музичких песама по појединим жанровима које су употребљене у конструкцији новог скупа података. Нека груба процена је шездесетак песама по жанру.

4. Анализа рада са кључним резултатима

Тема мастер рада дипл. инж. Драга Албијанића је препознавање музичког жанра песме употребом конволутивних неуралних мрежа. Музички жанрови као категорије су креирани од стране човека да би се одредила врста музике. Као и свака субјективна перцепција она је подложна категоријском преиспитивању. Суштина је да се рачунар преко метода машинског учења доведе до нивоа да у кратком времену обради велики број информација на начин који је близак човековом чулу, полазећи од неких основних звучних обележја музичке песме садржаних у форми слике на бази којих је могуће извести класификацију жанра коме она припада.

Овај рад се суштински састоји из две основне целине. Прве теоријске у којој је детаљно објашњена примена конволутивних неуралних мрежа у сврхе процене музичких жанрова, њихова структура, начин припреме аудио података првенствено применом разних функција из реда Фуријевих трансформација. Други део је имплементација једне веб апликације за препознавање музичких жанрова која користи знање стечено истраживањем теорије.

Циљ практичног дела овог рада је био да се састави и истренира модел који има способност да разликује музичке жанрове. Просечно људско биће песме успешно класификује у приближно 70% случајева. Оба модела која су истренирана достигла су ову стопу тачности. Први модел користи *GTZEN* скуп података и песме је сврстао у 10 различитих жанрова и показао се успешан у 73% случајева. Други модел користи новонастали скуп података који је доста временски обимнији од *GTZEN* скупа података. Број жанрова које овај модел класификује је 13. Тачност другог модела је 72%. Са порастом броја музичких жанрова опада и тачност модела пошто са тако расте могућност грешке при класификацији. Уколико би ова два модела имала исти број сличних жанрова постоји реална шанса да би новонастали скуп података дао бољи резултат.

4. Закључак и предлог

Кандидат Драго Албијанић је у свом мастер раду успешно решио проблем имплементације препознавања музичког жанра песме употребом конволутивних неуралних мрежа. Кандидат је исказао самосталност и систематичност у своме поступку као и иновативне елементе у решавању проблематике овог рада. У закључку рада дате су смернице у ком правцу би фокус наредних истраживања могао да иде. Паралелно са препознавањем музичког жанра интересантно је пре свега препознавања емоција које прате музичку песму.

На основу изложеног, Комисија предлаже Комисији за студије II степена Електротехничког факултета у Београду да рад дипл. инж. Драга Албијанића прихвати као мастер рад и кандидату одобри јавну усмену одбрану.

Београд, 01.09.2022. године

Чланови комисије:



Др Милош Цветановић, ванр.проф.



Др Саша Стојановић, доцент