



УНИВЕРЗИТЕТ У БЕОГРАДУ - ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ

Булевар краља Александра 73, 11000 Београд, Србија

Тел. 011/324-8464, Факс: 011/324-8681

КОМИСИЈИ ЗА СТУДИЈЕ II СТЕПЕНА ЕЛЕКТРОТЕХНИЧКОГ ФАКУЛТЕТА У БЕОГРАДУ

Комисија за студије II степена, Електротехничког факултета у Београду, на својој седници одржаној 08.06.2021. године именовало нас је у Комисију за преглед и оцену мастер рада дипл. инж. Милице Кланшчек под насловом „Генерисање аудио сигнала коришћењем *WaveNet* модела“. Након прегледа материјала Комисија подноси следећи

ИЗВЕШТАЈ

1. Биографски подаци кандидата

Милица Кланшчек је рођена 27.06.1996. године у Ваљеву. Завршила је основну школу „Сестре Илић“ у Ваљеву као вуковац. Уписала је Ваљевску гимназију коју је такође завршила као вуковац. Електротехнички факултет уписала је 2015. године. Дипломирала је на одсеку за Сигнале и системе 2019. године са просечном оценом 9,09. Дипломски рад одбранила је у септембру 2019. године са оценом 10. Дипломске академске – мастер студије на Електротехничком факултету у Београду, на Модулу за сигнале и системе уписала је у октобру 2019. године. Положила је све испите са просечном оценом 10.

2. Опис мастер рада

Мастер рад обухвата 84 стране, са укупно 60 слика, 4 табеле и 31 референцом. Рад садржи увод, 4 поглавља и закључак (укупно 6 поглавља), списак коришћене литературе, списак скраћеница, списак слика и списак табела.

Прво поглавље представља увод у коме су описани предмет и циљ рада. Укратко је представљен оригинални рад под називом „*WaveNet: A Generative Model for Raw Audio*“, који први пут представља *WaveNet* модел, основна идеја и мотивација иза рада. Дат је осврт на разлике између најчешће коришћених *TTS* (енгл. Text-To-Speech) система.

У другом поглављу је дат теоријски опис методе која је предмет овог рада. Представљене су теоријске основе потребне за разумевање методе. Описани су ауторегресивни модели, рекурентне и конволуционе неуралне мреже. Представљена су и два модела која претходе *WaveNet* моделу, а то су *PixelRNN* и *PixelCNN* модели. Детаљно су описани сви елементи који чине *WaveNet* модел, а на крају је описана и сама архитектура конволуционе неуралне мреже. Укратко су представљени резултати постигнути оригиналним радом.

У трећем поглављу је детаљно описана имплементација представљене методе. Описани су коришћени и анализирани скупови података, као и њихова обрада и припрема за модел. Описана је имплементирана архитектура *WaveNet* мреже, као и приступи коришћени приликом генерисања аудио сигнала.

Четврто поглавље детаљно описује имплементирани процес обучавања модела и представља различите структуре обучавањих модела, као и разлоге увођења истих. Структуре се међусобно разликују по параметрима саме архитектуре конволуционе неуралне мреже или структури и скупу коришћених података.

У оквиру петог поглавља је сваки од модела из четвртог поглавља описан понаособ, са додатним наведеним параметрима обучавања. За сваки од модела су приказани и анализирани резултати обучавања и каснијег генерисања аудио сигнала. Различите структуре и приступи процесу генерисања аудио сигнала су међусобно упоређени и објашњени су недостаци обучених модела, као и могућа даља побољшања.

Шесто поглавље је закључак у оквиру кога је сумиран приступ решавању дефинисаног проблема, добијени резултати, као и могућа даља унапређења представљене методе.

3. Анализа рада са кључним резултатима

Мастер рад дипл. инж. Милице Кланшчек се бави проблематиком генерисања аудио сигнала коришћењем *WaveNet* модела, са фокусом на музичке аудио сигнале. *WaveNet* модел, првобитно представљен у раду „*WaveNet: A Generative Model for Raw Audio*“ представља иновативан приступ моделовању аудио сигнала, коришћењем сировог сигнала. Овај модел је показао импресивне резултате, најпре у области генерисања говорног аудио сигнала, у *TTS* системима. Унапређена верзија *WaveNet* модела од 2017. године налази и примену у генерисању *Google Assistant* гласова, на северноамеричком енглеском и јапанском језику. У овом раду је оригинална метода најпре детаљно описана, а затим и имплементирана, након чега је модел обучаван коришћењем различитих параметара модела и скупова података. На крају су представљени и анализирани добијени резултати обучавања модела и генерисања аудио сигнала применом истих.

Основни доприноси рада су: 1) теоријски опис наведене методе за генерисање аудио сигнала; 2) имплементација и примена методе на јавно доступним скуповима аудио података, *IRMAS* и *VCTK*; 3) критичка анализа добијених резултата, 4) уочени могући правци за побољшање и даљи рад на овој теми.

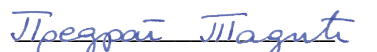
4. Закључак и предлог


Кандидат Милица Кланшчек је у свом мастер раду описала, имплементирала и применила метод машинског учења у циљу генерисања аудио сигнала, познатог као *WaveNet* модел. Критички је анализирали и упоредили резултате добијене коришћењем различитих изабраних параметара модела и података за обучавање истог, као и приступа генерисању аудио сигнала. Предложила је могуће приступе у даљем истраживању. Исказала је самосталност и систематичност у свом поступку.

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

Београд, 03.09.2021. године

Чланови комисије:


Др Предраг Тадић, доцент


Др Александра Марјановић, доцент