

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки
кафедра ЕОМ

ДОСЛІДЖЕННЯ ПРОДУКТИВНОСТІ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ ПРИ СЕМАНТИЧНОМУ АНАЛІЗІ

Кваліфікаційна робота
Другий рівень (магістр)

Автор

Литвиненко В. С.
ст. гр. СМм-20-2

Керівник

Фесенко Т.Г.
проф. каф. ЕОМ

МЕТА КВАЛІФІКАЦІЙНОЇ РОБОТИ

Метою роботи є дослідження продуктивності різних нейромережевих моделей при семантичному аналізі аудіо записів, а також реферування тексту.

ЗАДАЧІ КВАЛІФІКАЦІЙНОЇ РОБОТИ

- розглянути і порівняти існуючі моделі для перетворення мовлення в текст;
- розглянути і порівняти існуючі моделі для сумаризації тексту;
- забезпечити можливість працювати вище переліченим моделям з українською мовою;
- оцінити якість роботи даних моделей з українською та англійською мовами;
- оцінка таймінгу роботи розробленої системи;
- аналіз отриманих результатів.

3

АКТУАЛЬНІСТЬ КВАЛІФІКАЦІЙНОЇ РОБОТИ

- збільшення обсягів інформації у цифровому вигляді;
- зростання вимог щодо швидкості обробки та доступу до інформації;
- підвищення продуктивності в пошуку інформації;
- отримання анотації (резюме, реферату);
- широкий спектр областей застосування.

Сфери застосування:

Лекції

Виступи на конференція

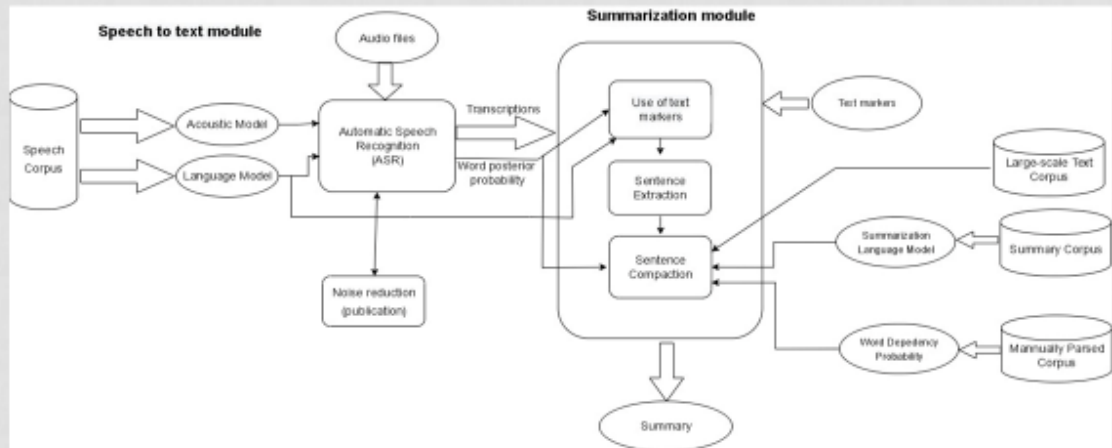
Захист наукових робіт

4

ОГЛЯД ІСНУЮЧИХ АНАЛОГІВ

Критерії	Програмні рішення		
	Amazon Transcribe	Dragon Anywhere	QuillBot
Підтримка обробки голосу в текст	+	+	-
Підтримка функціоналу сумаризації	-	-	+
Функціонал працює з українською мовою	+	-	-
Функціонал працює з англійською мовою	+	+	+
Варіативність режимів роботи	-	-	-

РІШЕННЯ ПОСТАВЛЕНОЇ ЗАДАЧІ. АРХІТЕКТУРА ЗАПРОПОНОВАНОЇ МОДЕЛІ



АПАРАТНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ТЕСТУВАННЯ

CPU	Intel Core i7-9750H (2.6 - 4.5 ГГц)	Intel Xeon 2.30GHz
GPU	NVIDIA GeForce GTX 1650 Mobile	NVIDIA Tesla T4
GPU (характеристики)	896 ядер CUDA, 1665 МГц, 4 Гб DDR5	2560 ядер CUDA, 320 ядер Tensor, 1250 МГц, 16 Гб DDR6
RAM	16 Gb	12 Gb
OS	Linux (Ubuntu 20.04)	Linux

7

АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SPEECH TO TEXT. ЕТАП ОЧИЩЕННЯ ВІД СТОРОНЬОГО ШУМУ

Ефективність очищення від стороннього шуму

Мова	Українська	Англійська
Аудіодоріжка очищена від шуму, WER	29	3.9
Аудіодоріжка не очищена від шуму, WER	38.7	5.2
Різниця, %	25%	21%

8

АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SPEECH TO TEXT. ЕТАП ОТРИМАННЯ ТЕКСТУ З АУДІО

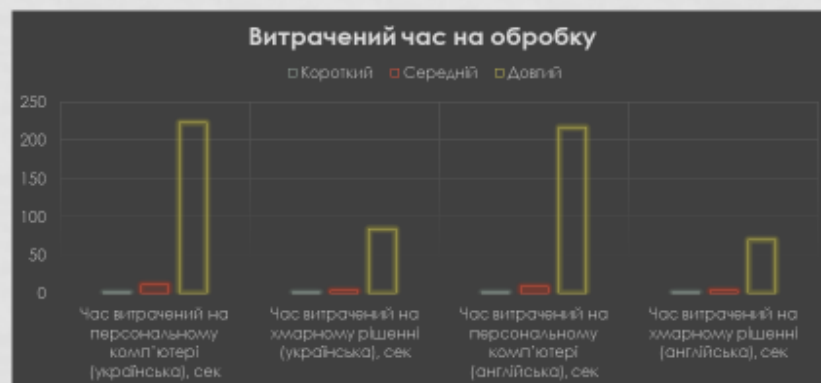
Довжина аудіозапису	Час витрачений на персональному комп'ютері (українська), сек	Час витрачений на хмарному рішенні (українська), сек	Час витрачений на персональному комп'ютері (англійська), сек	Час витрачений на хмарному рішенні (англійська), сек
Короткий	0.8	0.6	0.6	0.48
Середній	10.8	3.6	9.4	3.14
Довгий	222.7	84.2	217.1	70.1

Класифікація записів:

- a) короткий (5с)
- b) середній (~ 60с)
- c) довгий (~ 240с).

9

АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SPEECH TO TEXT. ЕТАП ОТРИМАННЯ ТЕКСТУ З АУДІО



10

АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SPEECH TO TEXT. ЕТАП ОТРИМАННЯ ТЕКСТУ З АУДІО

Довжина аудіозапису	Українська мова, WER	Англійська мова, WER
Короткий	24.6	3.4
Середній	28.4	3.7
Довгий	33.9	4.5



11

АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SUMMARIZATION. УКРАЇНСЬКА МОВА

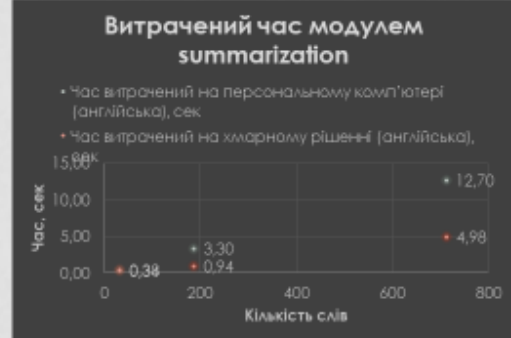
Кількість слів	Час витрачений на персональному комп'ютері (українська), сек	Час витрачений на хмарному рішенні (українська), сек
25	0.4	0.37
136	3.2	0.92
725	13.6	5.13



12

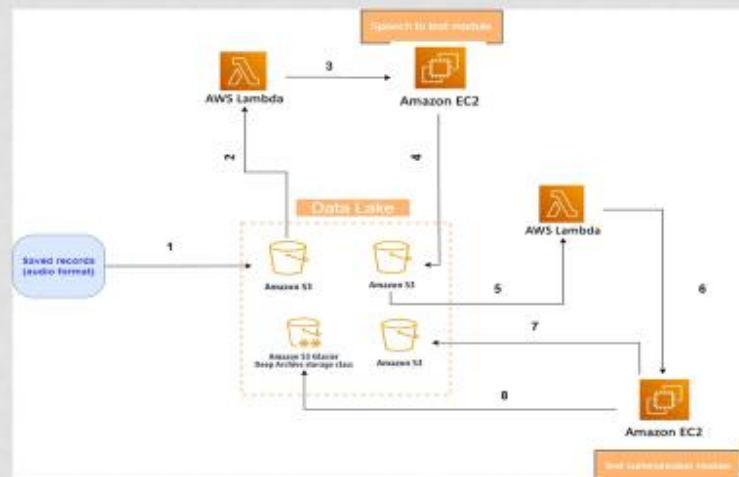
АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ МОДУЛЬ SUMMARIZATION. УКРАЇНСЬКА МОВА

Кількість слів	Час витрачений на персональному комп'ютері (англійська), сек	Час витрачений на хмарному рішенні (англійська), сек
32	0.38	0.34
187	3.3	0.94
713	12.7	4.98



13

АРХІТЕКТУРА CLOUD РІШЕННЯ



14

Стаття у фаховому виданні:

Barkovska O., Kholiev, V., & Lytvynenko, V. (2022). STUDY OF NOISE REDUCTION METHODS IN THE SOUND SEQUENCE WHEN SOLVING THE SPEECH-TO-TEXT PROBLEM. *Advanced Information Systems*, 6(1). 48–54. <https://doi.org/10.20998/2522-9052.2022.1.08>

Тези доповіді:

Барковська О. Ю., Литвиненко В. С. Дослідження продуктивності нейромережових моделей при семантичному аналізі / Сучасні напрями розвитку Інформаційно-комунікаційних технологій та засобів управління. Тези доповідей 12-ої МНТК (online-режим). – 2022. – Т.1: секція 3. – с.136.



15

ВИСНОВКИ

- Розпізнавання мовлення являється придатним і відносно надійним інструментом для розпізнавання голосу з аудіофайлів. Але для того, щоб підвищити якість і точність потрібна попередня обробка аудіозапису, а також натренована модель на лінгвістичному корпусі.
- Сумаризація тексту являється хорошим інструментом для узагальнення інформації. Але має обмеження щодо втрати контексту, якщо реферується великий текст. Запропонована ідея використання слів маркерів направлена на часткове вирішення цієї проблеми.
- В процесі виконання кваліфікаційної роботи було проведено дослідження параметрів, які повинна задовольняти система розпізнавання голосу та система сумаризації тексту. Також були розглянуті існуючі методи обробки NLP у випадку з перетворенням мовлення в текст і його сумаризації.

16