

# РОЗРОБКА МЕТОДУ РОЗПІЗНАВАННЯ УКРАЇНСЬКОГО МОВЛЕННЯ МЕДИЧНОГО СПРЯМУВАННЯ З ПЕРЕТВОРЕННЯМ АУДІОЗАПИСІВ У ТЕКСТ

Вінницький національний технічний університет

## *Анотація*

*У роботі здійснено детальний аналіз методів розпізнавання українського медичного мовлення та перетворення аудіозаписів у текст. Дослідження охоплює визначення ключових понять, впливаючих на процес розпізнавання, та розробку технічного завдання для створення нового програмного засобу. Представлено архітектуру програмного забезпечення, включаючи графічний інтерфейс та модулі обробки мовлення, з використанням Python та сучасних технологій. Розкрито методику автоматизації збору датасету, що включає понад 2000 унікальних аудіозаписів, та описано високу точність розпізнавання медичних текстів за допомогою реалізованої системи з використанням моделі Whisper та техніки LoRA. Ефективність системи підтверджено через експериментальну апробацію та використання метрики WER.*

**Ключові слова:** розпізнавання українського медичного мовлення, перетворення аудіозапису у текст, обробка природної мови, розробка програмного забезпечення, автоматизація збору датасету, LoRA, Whisper, WER.

## *Abstract*

*This research work conducts a detailed analysis of methods for recognizing Ukrainian medical speech and converting audio recordings into text. The study encompasses the identification of key concepts influencing the speech recognition process and the development of a technical task for creating a new software tool. The architecture of the software, including the graphical interface and speech processing modules, using Python and modern technologies, is presented. The methodology for automating the dataset collection, including over 2000 unique audio recordings, is revealed, and the high accuracy of medical text recognition using the implemented system with the Whisper model and LoRA technique is described. The system's effectiveness is confirmed through experimental testing and the use of the WER metric.*

**Keywords:** recognition of Ukrainian medical speech, conversion of audio recording into text, natural language processing, software development, dataset collection automation, LoRA, Whisper, WER.

## **Вступ**

У епоху швидкого розвитку інформаційних технологій, де швидкість та точність обробки даних є ключовими для успіху у багатьох сферах, зокрема у медичній, набуває важливості розвиток інноваційних технологій.

Особливо важливим стає питання ефективного використання цих технологій для підвищення якості медичного обслуговування, зокрема через автоматизацію документації та аналізу мовленнєвої інформації. Саме тому актуальним є розробка методів розпізнавання українського медичного мовлення та їх перетворення у текст [1].

Ця робота покликана вирішити проблему підвищення ефективності медичного обслуговування шляхом автоматизації процесів розпізнавання та обробки мовлення. Вона спрямована на розробку методу автоматизованого розпізнавання українського медичного мовлення та його перетворення на текстовий формат. Проведено аналіз існуючих технологій, визначено особливості українського медичного мовлення, розроблено алгоритми та методи для розпізнавання мовлення, що враховують ці особливості.

Практичною цінністю роботи є можливість її застосування у медичній практиці для оптимізації процесу документування та обробки мовленнєвої інформації, що сприятиме поліпшенню якості медичного обслуговування. Апробація розробленої методики показала її ефективність та підтвердила наукову новизну одержаних результатів.

## Результати дослідження

У ході дослідження для реалізації поставленої задачі автоматизованого розпізнавання українського медичного мовлення було обрано мову програмування Python. Цей вибір зумовлений гнучкістю та великим набором бібліотек, які полегшують розробку складних систем обробки природної мови.

Важливою частиною процесу стала розробка архітектури системи (рис. 1), орієнтованої на збір та обробку даних.

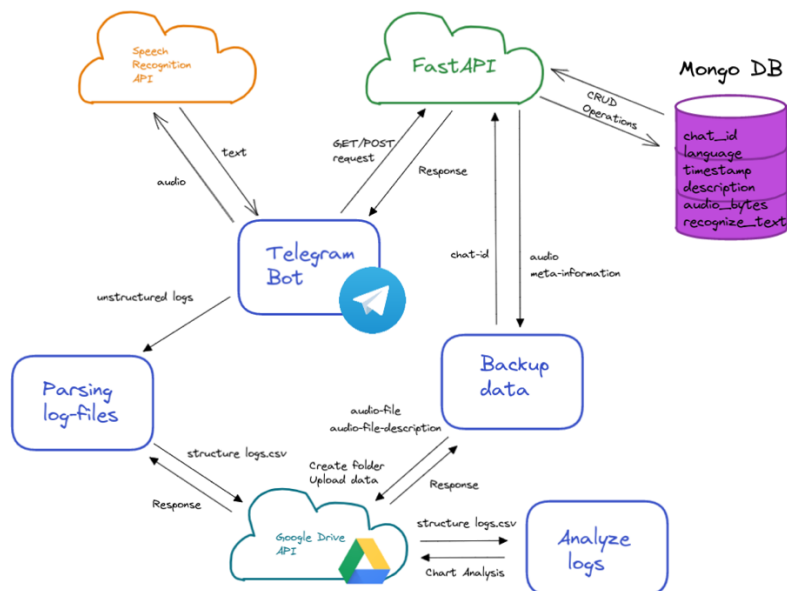


Рис. 1. Схема архітектури сервісів

Основною інновацією у цій частині стала інтеграція модулів для автоматизованого анотування медичних текстів. Це дозволило створити структуровану базу даних, яка містить унікальний набір аудіозаписів медичного мовлення, зібраний з екзаменів "КРОК" [2]. Такий підхід забезпечив наявність високоякісного та релевантного датасету для подальшого тренування моделі.

Значний акцент у дослідженні було зроблено на тренуванні моделі Whisper для розпізнавання українського мовлення. Завдяки адаптації моделі з використанням техніки LoRA (Low-Rank Adaptation), було можливо забезпечити високу точність та ефективність розпізнавання [3, 4]. Незважаючи на складність медичної термінології та особливості української мови, система показала обнадійливі результати, відкриваючи перспективи для подальшого вдосконалення та розширення застосування моделі в медичній галузі.

Розроблена програмна система представляє собою важливий крок у напрямку покращення обробки медичного мовлення, що є ключовим для підвищення якості медичних послуг та оптимізації роботи медичних фахівців. Зібрані дані та отримані результати демонструють значний потенціал для майбутніх досліджень у цій важливій області.

## Висновки

В ході проведеного дослідження було здійснено аналіз існуючих методів розпізнавання українського медичного мовлення та його перетворення в текст, з акцентом на розробку програмного засобу для цієї мети. Дослідження включало аналіз ключових понять, що впливають на процес розпізнавання мовлення, та формування технічного завдання для розробки нової системи.

Значна увага в роботі була приділена розробці архітектури програмного забезпечення, що охоплює як графічний інтерфейс, так і модулі обробки мовлення. Використання Python та інших сучасних технологій сприяло створенню ефективної, інтуїтивно зрозумілої та надійної системи.

Основним досягненням стало розроблення методу автоматизації збору датасету, який зібрав значну кількість унікальних аудіозаписів українського медичного мовлення, відкриваючи нові можливості для дослідження та аналізу медичного спілкування [5]. Застосування моделі Whisper та техніки LoRA

у тренуванні моделі продемонструвало високу точність розпізнавання, що робить цю систему перспективною для широкого використання у медичній галузі.

У результаті, розроблена програмна система не лише відповідає сучасним вимогам до ефективності та точності обробки мовлення, але й має значний потенціал для подальшого розвитку та вдосконалення, що відкриває широкі перспективи для її застосування у медичній практиці.

#### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Бісікало О. Огляд аналогів програм та технологій для розпізнавання українського мовлення з перетворенням у текст / О.В. Бісікало, П.О. Петрук // Актуальні задачі медичної, біологічної фізики та інформатики : мат. статей Всеукр. наук.-практ. конф. з міжн. участю (27 квітня 2022 р., ВНМУ ім. М.І. Пирогова, м. Вінниця). – Вінниця : ВНМУ ім. М.І. Пирогова, 2022. – С.79-82.
2. Крок - Підготовка до тестування та екзаменів. – Режим доступу: <https://testkrok.org.ua/>.
3. GitHub. – Режим доступу: <https://github.com/openai/whisper>.
4. Exploring efficient-tuning methods in self-supervised speech models – Режим доступу: <https://arxiv.org/pdf/2210.06175.pdf>.
5. Telegram – Режим доступу: [https://t.me/Speech\\_to\\_Text\\_test\\_version\\_Bot](https://t.me/Speech_to_Text_test_version_Bot).

**Бісікало Олег Володимирович** — д-р техн. наук, професор, завідувач кафедри АІТ, Вінницький національний технічний університет, м. Вінниця e-mail: [obisikalo@vntu.edu.ua](mailto:obisikalo@vntu.edu.ua)

**Петрук Петро Олександрович** — студент групи ЗАКІТ-22м, факультет автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: [petr.square@gmail.com](mailto:petr.square@gmail.com)

**Bisikalo Oleg V.** — Dr.Sc. (Eng.), Professor of the Faculty for Computer Systems and Automatic, Vinnytsia National Technical University, Vinnytsia

**Petruk Petro O.** — Department of Automation and intelligent information technologies, Vinnytsia National Technical University, Vinnytsia, email : [petr.square@gmail.com](mailto:petr.square@gmail.com)