

УДК 004.934.2

Гладиш Д. – ст. гр. СП-41

*Тернопільський національний технічний університет імені Івана Пулюя*

## **РОЗРОБКА СИСТЕМИ ГОЛОСОВОГО ВВЕДЕННЯ ТЕКСТУ В АКТИВНЕ ПОЛЕ УКРАЇНСЬКОЮ МОВОЮ З ПІДТРИМКОЮ ОРФОГРАФІЧНОГО АНАЛІЗУ**

Науковий керівник: к.т.н., ст. викл. Тимків П. О.

Hladysh D. O.

*Ternopil Ivan Puluj National Technical University*

## **DEVELOPMENT OF A UKRAINIAN VOICE TEXT INPUT SYSTEM FOR ACTIVE FIELDS WITH SPELL-CHECK SUPPORT**

Supervisor: P. O. Tymkiv, PhD in Engineering, Senior Lecturer

Ключові слова: голосове введення, розпізнавання мовлення, українська мова, Whisper, офлайн-система.

Keywords: voice input, speech recognition, Ukrainian language, Whisper, offline system.

Сучасні засоби голосового введення тексту є важливим напрямом розвитку людино-машинної взаємодії, оскільки дають змогу спростити введення інформації, підвищити швидкість роботи користувача та розширити доступність програмних систем. Особливо актуальним є створення таких рішень для української мови, оскільки наявні інструменти не завжди забезпечують достатній рівень автономності, точності та зручності інтеграції з настільними програмами. У зв'язку з цим доцільною є розробка програмної системи голосового введення тексту в активне поле з орієнтацією на локальну роботу без використання віддалених сервісів.

Метою роботи є розробка desktop-системи голосового введення тексту українською мовою, що працює в повністю офлайн-режимі, підтримує інтеграцію з активним полем вводу в довільних застосунках та забезпечує орфографічний аналіз розпізаного тексту. Основне призначення системи полягає в забезпеченні альтернативного способу введення текстової інформації без використання клавіатури – шляхом активації гарячою клавішею, голосового введення та автоматичної вставки обробленого тексту.

Програмна реалізація системи виконується мовою Python 3.11. Архітектура проекту є модульною і складається з окремих компонентів: модуля запису аудіо, рушія розпізнавання, модуля вставки тексту, менеджера гарячих клавіш, бази даних історії та графічного інтерфейсу. Для побудови графічного інтерфейсу використовується стандартна бібліотека tkinter, для запису аудіо – PyAudio та numpy, для локального офлайн-розпізнавання мовлення – модель OpenAI Whisper з апаратним прискоренням через CUDA. Вставка розпізаного тексту в активне поле реалізована через бібліотеки pyperclip та pyautogui, що забезпечує сумісність з довільними застосунками на платформі Windows.

На поточному етапі реалізовано повноцінне ядро системи, яке забезпечує повний цикл обробки мовлення: запис аудіо з мікрофона, локальне розпізнавання українського мовлення та автоматичну вставку результату в активне поле.

Функціонування системи організовано у вигляді послідовного pipeline: запис аудіо у форматі WAV, передача аудіоданих моделі Whisper, отримання текстового результату, орфографічна перевірка через LanguageTool та вставка тексту в активне поле через буфер обміну. Керування записом здійснюється за допомогою глобальної гарячої клавіші, що реалізована через бібліотеку rnpur і працює у фоновому режимі незалежно від активного застосунку.

Ключовою перевагою запропонованого підходу є повністю офлайн-режим роботи. Це дозволяє зменшити залежність від мережевого з'єднання, підвищити автономність системи та забезпечити можливість використання в умовах, де доступ до Інтернету є обмеженим або небажаним. Додатковою перевагою є універсальність інтеграції, оскільки система не обмежується одним конкретним середовищем, а може використовуватися в текстових редакторах, полях введення, формах і месенджерах.

Проведене тестування підтвердило працездатність розробленого рішення. Перевірено запис аудіо з мікрофона, коректність формування WAV-файлів, роботу моделі Whisper medium з апаратним прискоренням на GPU NVIDIA RTX 3060, розпізнавання української мови та вставку тексту в активне поле різних застосунків – браузера, текстового редактора та середовища розробки. Додатково перевірено роботу орфографічного аналізу в режимах автоматичного виправлення та ручного підтвердження. У результаті встановлено, що система функціонує стабільно, час розпізнавання 10-секундного фрагменту становить 1–2 секунди, а точність розпізнавання української мови є достатньою для практичного використання.

Практична цінність роботи полягає у створенні інструмента для швидкого голосового введення українськомовного тексту в довільних прикладних програмах без використання інтернету. Система реалізує повний цикл: від активації гарячою клавішею – через розпізнавання мовлення та орфографічну перевірку – до автоматичної вставки тексту в активне поле. Розроблений графічний інтерфейс надає доступ до історії розпізнаних записів, редактора тексту перед вставкою та гнучких налаштувань системи. Рішення може бути корисним для широкого кола користувачів, зокрема для тих, хто прагне прискорити введення тексту, а також для осіб з обмеженими можливостями роботи з клавіатурою.