«Mnemozia»

Telegram-бот для транскрибации аудиосообщений в текст.

Образовательная организация: ГБПОУ “СПК”
Автор:Никифоров Артём Владимирович
Руководитель:Кувшинова Светлана Михайловна

Проблемы, которые решает проект, и идея проекта
Проблема: современные коммуникации перегружены голосовыми сообщениями, которые неудобно или невозможно слушать в учебе, на работе или в публичных местах.
Решение – перевод их в текст с помощью нейросети

Идея: Разработка Telegram-бота, на основе нейронной искусственной сети “**OpenAi Whisper**”.
для быстрого и конфиденциального перевода голосовых сообщений в текстовый формат, без посредников
и ограничений

Цель проекта

Цель: Создать бесплатный локальный инструмент для быстрой расшифровки аудио с удобным интерфейсом и без ограничений по длине

Задачи:

1. Интегрировать нейросеть
2. Настроить GPU-ускорение
3. Написать интерфейс Telegram-бота

Технические подробности проекта

Язык программирования:
Python
Ядро ИИ:
faster-whisper large-v3
Платформа: NVIDIA CUDA
Встроенные модули Python:
asyncio, os, sys, time, logging, pathlib, site

Со стороны Telegram:
aiogram для скачивания в буфер аудиофайлов.
Пакетная обработка данных
Полная локальность(Запуск производится на ПК)
Данные хранятся только на момент обработки

Практическая демонстрация и апробация

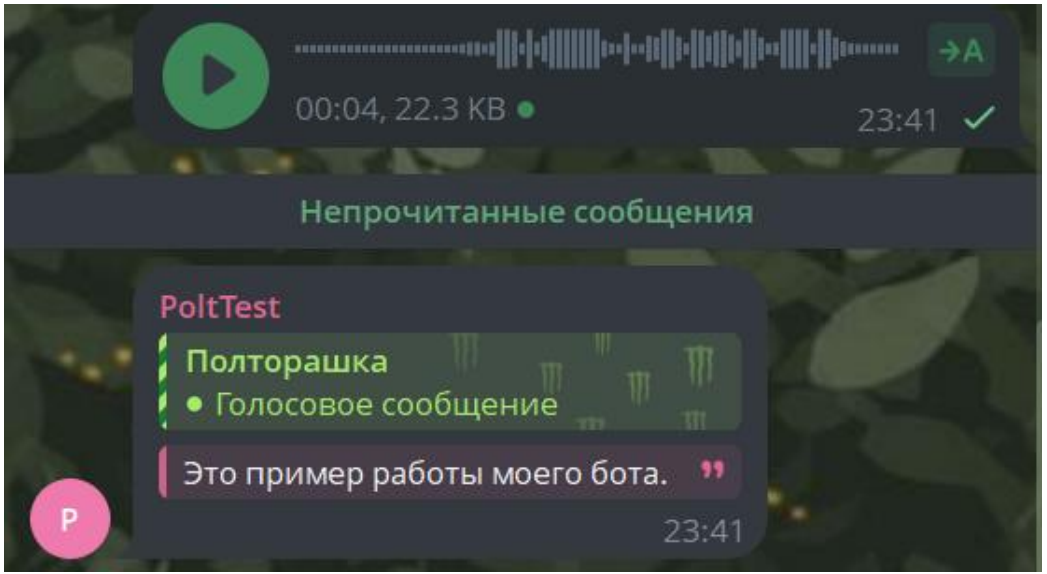


Рисунок 1 – пример работы Telegram-бота «Mnemozia»

Бот успешно прошел этап практического тестирования в студенческой среде, было протестирована расшифровка десятков аудиосообщений разной длинны
Подтверждена стабильность работы

Сравнение изменения скорости транскрибации

Таблица 1. Сравнение производительности

Оборудование	Длительность аудио	Время обработки	Ускорение
CPU: Intel Core i5-12400F	65.2 сек.	55.24 сек.	~1.0x
GPU: NVIDIA RTX 4060	65.2 сек.	7.14 сек	~7.7x

Практическая ценность, и результаты проекта

Для достижения цели был разработан полностью автономный Telegram-бот на Python
Результат:
Создан стабильно работающий инструмент Ускоряющий расшифровку и помогающий в быт

Проект решает проблему мессенджеров длинных голосовых сообщений экономя время.
Конфиденциальность: обработка проходит локально, данные хранятся у пользователя

Заключение и планы развития

Итоги проекта: Разработан и полностью готов к работе автономный Telegram-бот «Mnemozia» для локальной транскрибации аудио

Также планируется внедрение диаризации спикеров: разделение аудиопотока по голосам и автоматическое форматирование расшифровки в виде понятного диалога (кто именно и что сказал)