

Инструмент анализа голосов на аудиозаписи

Лидер, студент гр. 6304

Пискунов Ярослав Андреевич

Команда, студенты гр. 8303

Быков Андрей Владимирович

Деркач Никита Владимирович

Логинов Евгений Анатольевич

Суть проекта

Цель: создать приложение командной строки, которое будет принимать на вход аудиофайлы (mp3) и обрабатывать их.

Функционал:

- идентификация диктора по набору контрольных образцов,
- идентификация наличия нескольких дикторов на аудио,
- вычисление и вывод фрагментов аудиозаписи для каждого из дикторов в файлы,
- распознавание речи для каждого диктора (с указанием меток времени) с выводом в файл.

Задачи на четвертую итерацию

- объединение всего разработанного функционала в единую систему;
- разработка и внедрение метрики уверенности в результате;
- поиск и устранение ошибок;
- калибровка точности распознавания.

Метрика уверенности в результате

Для каждой реплики в файле выводится степень уверенности в точности распознавания

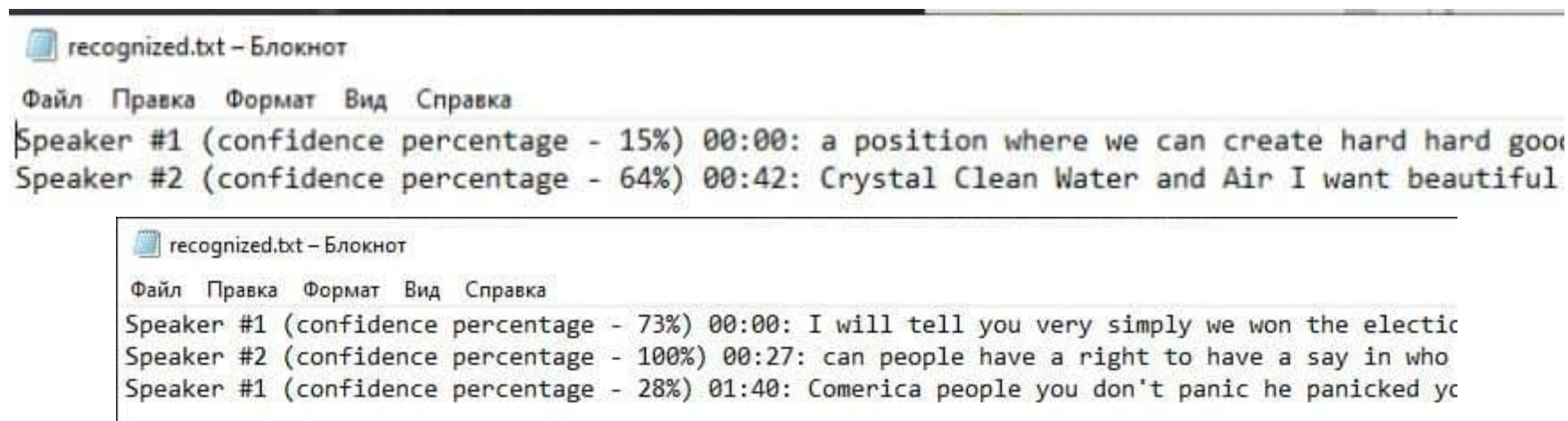


Рисунок 1 – Точность распознавания в итоговом файле

Достигнутый результат

- ✓ разработанный ранее функционал объединен в единое приложение;
- ✓ разработана и внедрена метрика уверенности в результате
- ✓ исправлена ошибка, не позволявшая приложению корректно работать в Docker-контейнере, а также ряд других ошибок, связанных функциональностью;
- ✓ точность распознавания откалибрована, чтобы результаты были как можно более точными.

Репозиторий с проектом - https://github.com/moevm/mse_voice_analysis

Достигнутый результат. Юнит-тесты

Здесь будут скрины юнит тестов

Рисунок 2 – Тесты

Звуковые файлы, использованные в тестах

https://github.com/moevm/mse_voice_analysis/tree/debug/res/audio_data_test

Достигнутый результат. Демонстрация

Здесь будет скринкаст

Достигнутый результат. Демонстрация 2

Здесь будет другой скринкаст

Пути дальнейшего развития проекта

- разработка графического интерфейса для приложения;
- улучшение и/или изменение алгоритмов распознавания для повышения точности и уверенности в результате;
- оптимизация работы приложения с целью уменьшения времени обработки звуковых файлов;
- поиск и исправление ошибок.

Q & A