

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 третью итерацию

- разработка функционала вычисления и вывода фрагментов аудиозаписи для каждого из дикторов в файлы;
- разработка функционала распознавания речи для каждого диктора (с указанием меток времени) с выводом в файл;
- написание юнит-тестов для указанного выше функционала с использованием собранных звуковых файлов;
- расширение количества юнит-тестов для разработанного на второй итерации функционала за счет собранных звуковых файлов;
- приведение функционала в соответствие с макетом и ожидаемым результатом.

Макет

```
yp@yp-VirtualBox:~/Documents/mse_voice_analysis$ mse_voice_analysis ../../Downloads/Tony_Igy_-_Astronomia_14893.mp3 --distribution
Processing file
Speech recognition complete. See "results" folder for splited fragments for each speaker
```

Рисунок 1 – Вычисление и вывод фрагментов для каждого из дикторов в файлы

```
yp@yp-VirtualBox:~/Documents/mse_voice_analysis$ mse_voice_analysis ../../Downloads/Tony_Igy_-_Astronomia_14893.mp3 --speech
Processing file
Speech recognition for each speaker complete. See "results" folder for timestamps
```

Рисунок 2 – Распознавание речи для каждого из дикторов

Методы и технологии

- распознавание речи производится с использованием библиотеки `speech_recognition`, которая позволяет без проблем выполнить перевод из голоса в текст;
- разбиение звукового файла на отдельные звуковые фрагменты осуществляется с использованием `AudioSegment` из библиотеки `pydub`.

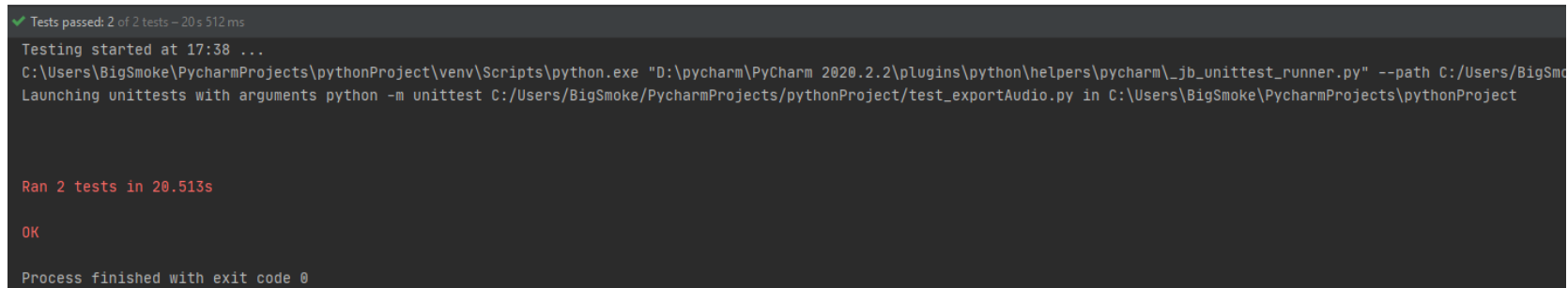
Достигнутый результат

- ✓ разработан функционал вычисления и вывода фрагментов аудиозаписи для каждого из дикторов в файлы;
- ✓ разработан функционал распознавания речи для каждого диктора;
- ✓ для разработанного функционала созданы юнит-тесты;
- ✓ расширено количество юнит-тестов для функционала, разработанного на прошлой итерации;
- × приведение функционала в соответствие с макетом и ожидаемым результатом.

Создан Docker-контейнер для приложения.

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

Достигнутый результат. Юнит-тесты



```
✓ Tests passed: 2 of 2 tests - 20 s 512 ms
Testing started at 17:38 ...
C:\Users\BigSmoke\PycharmProjects\pythonProject\venv\Scripts\python.exe "D:\pycharm\PyCharm 2020.2.2\plugins\python\helpers\pycharm\_jb_unittest_runner.py" --path C:/Users/BigSmoke\PycharmProjects\pythonProject
Launching unittests with arguments python -m unittest C:/Users/BigSmoke/PycharmProjects/pythonProject/test_exportAudio.py in C:/Users/BigSmoke/PycharmProjects/pythonProject

Ran 2 tests in 20.513s

OK

Process finished with exit code 0
```

Рисунок 3 – Тесты вычисления и вывода фрагментов аудиозаписи в файл



```
✓ Tests passed: 4 of 4 tests - 4 s 9 ms
Recognition status: 1/1...
Ran 4 tests in 4.012s

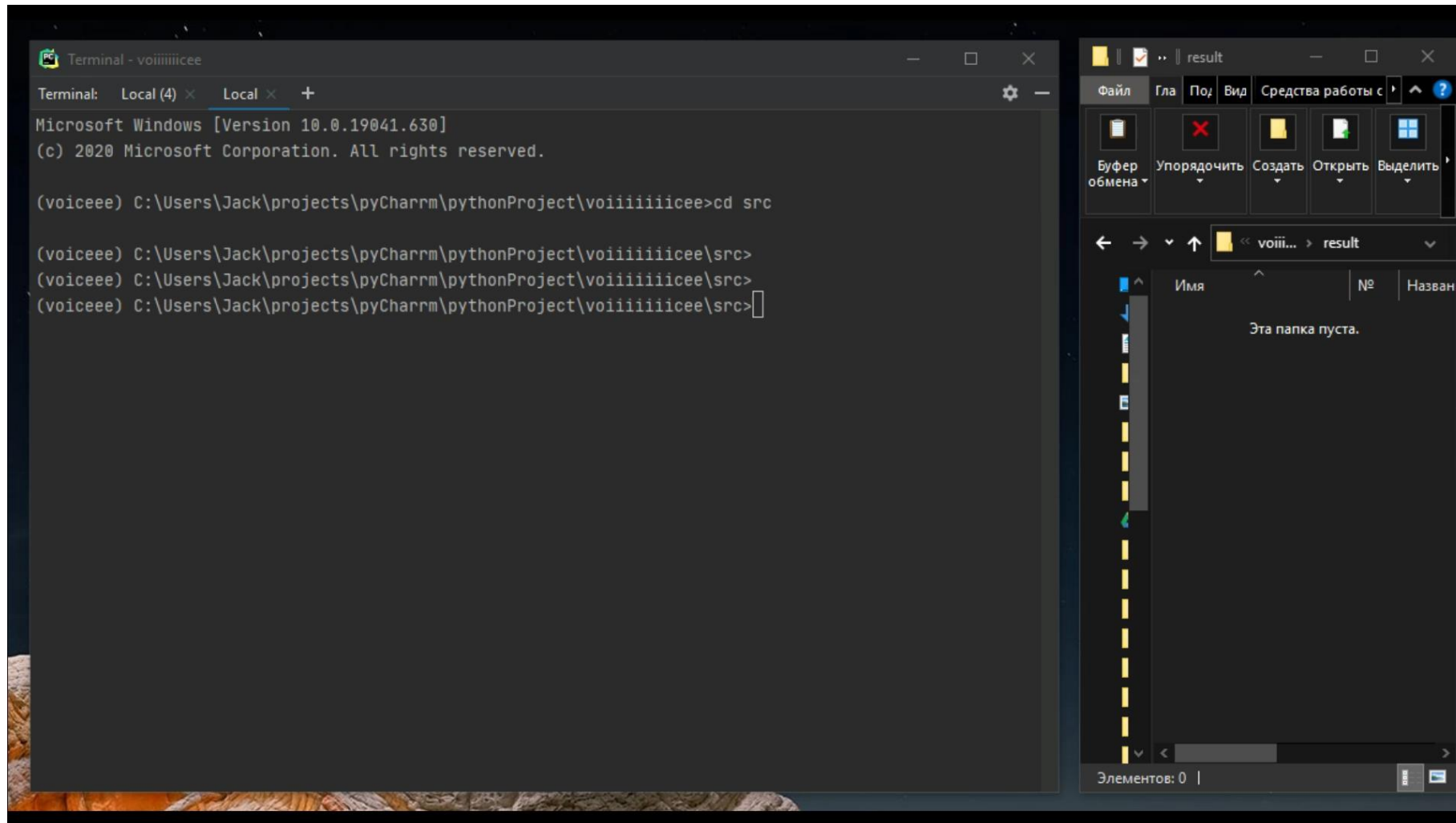
OK
```

Рисунок 4 – Тесты распознавания речи для каждого диктора

Звуковые файлы, использованные в тестах

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

Достигнутый результат. Демонстрация



Задачи на четвертую итерацию

- объединение всего разработанного функционала в единую систему;
- приведение функционала в соответствие с ожидаемым результатом и макетом;
- поиск и устранение ошибок, расширение количества юнит-тестов.

Подробнее о задачах -

https://github.com/moevm/mse_voice_analysis/projects

Q & A