

# 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

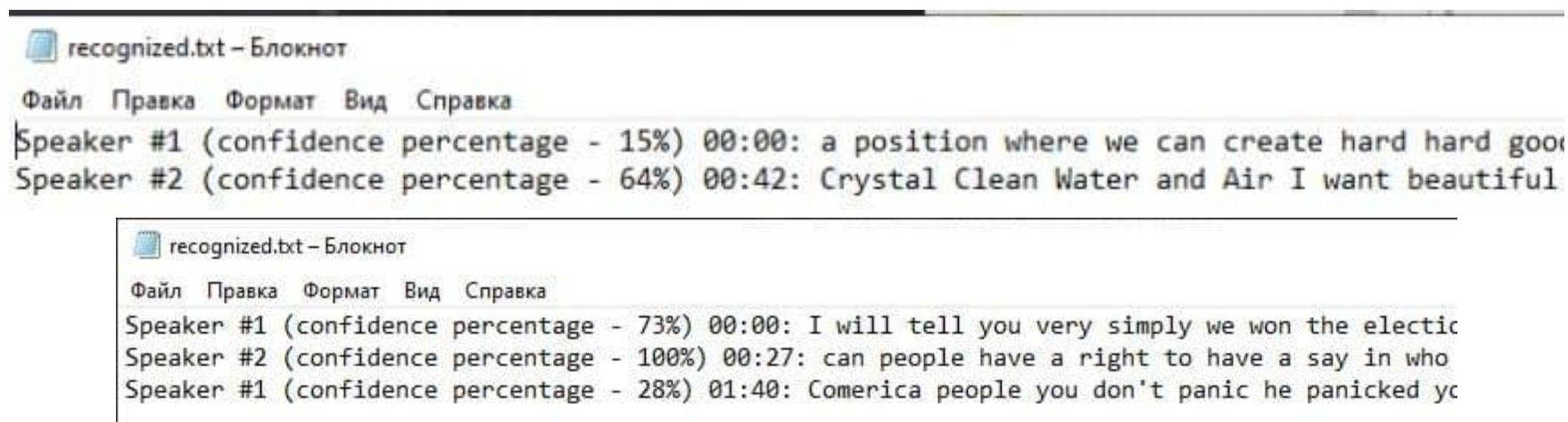


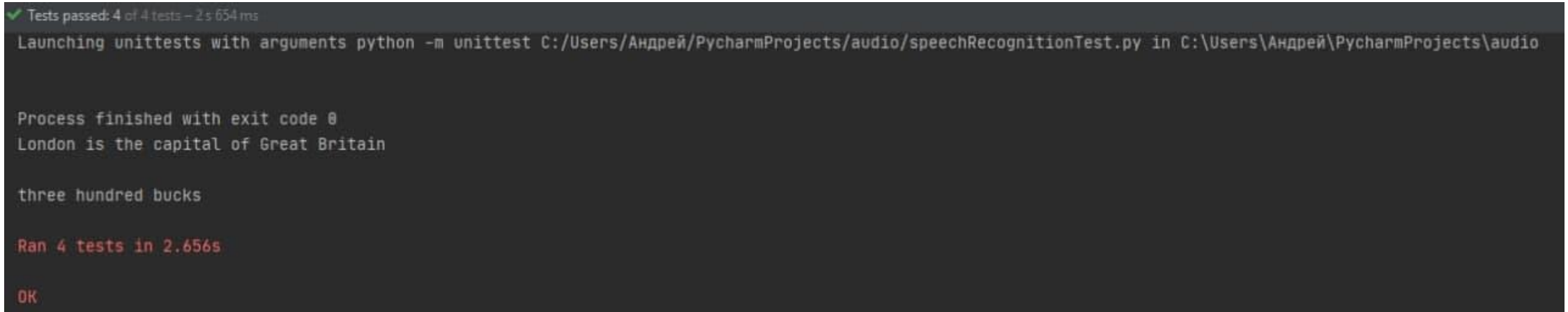
Рисунок 1 – Точность распознавания в итоговом файле

# 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](https://github.com/moevm/mse_voice_analysis)

# Достигнутый результат. Юнит-тесты



```
✓ Tests passed: 4 of 4 tests – 2 s 654 ms
Launching unittests with arguments python -m unittest C:/Users/Андрей/PycharmProjects/audio/speechRecognitionTest.py in C:\Users\Андрей\PycharmProjects\audio

Process finished with exit code 0
London is the capital of Great Britain

three hundred bucks

Ran 4 tests in 2.656s

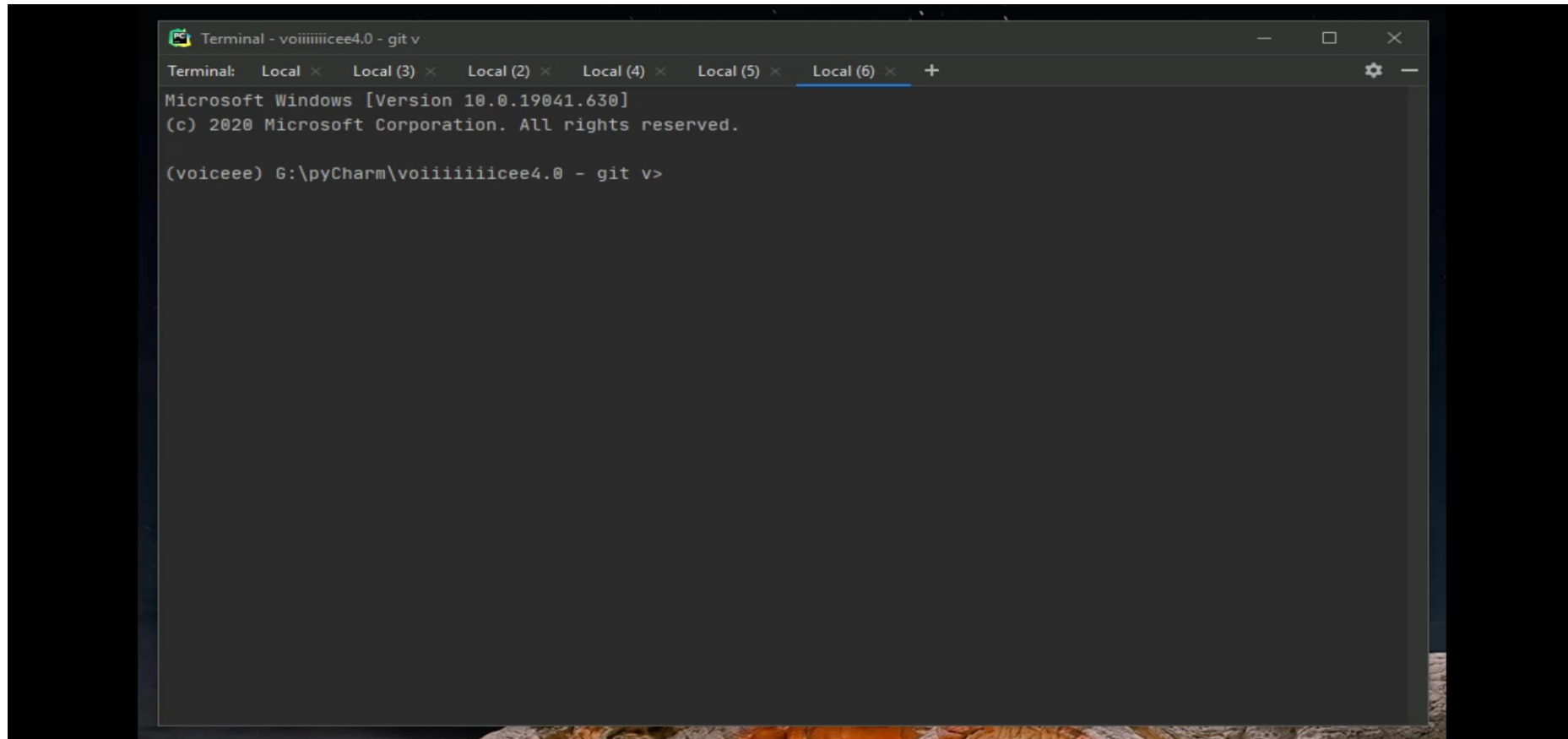
OK
```

Рисунок 2 – Тесты

Звуковые файлы, использованные в тестах

[https://github.com/moevm/mse\\_voice\\_analysis/tree/debug/res/audio\\_data\\_test](https://github.com/moevm/mse_voice_analysis/tree/debug/res/audio_data_test)

# Достигнутый результат. Демонстрация

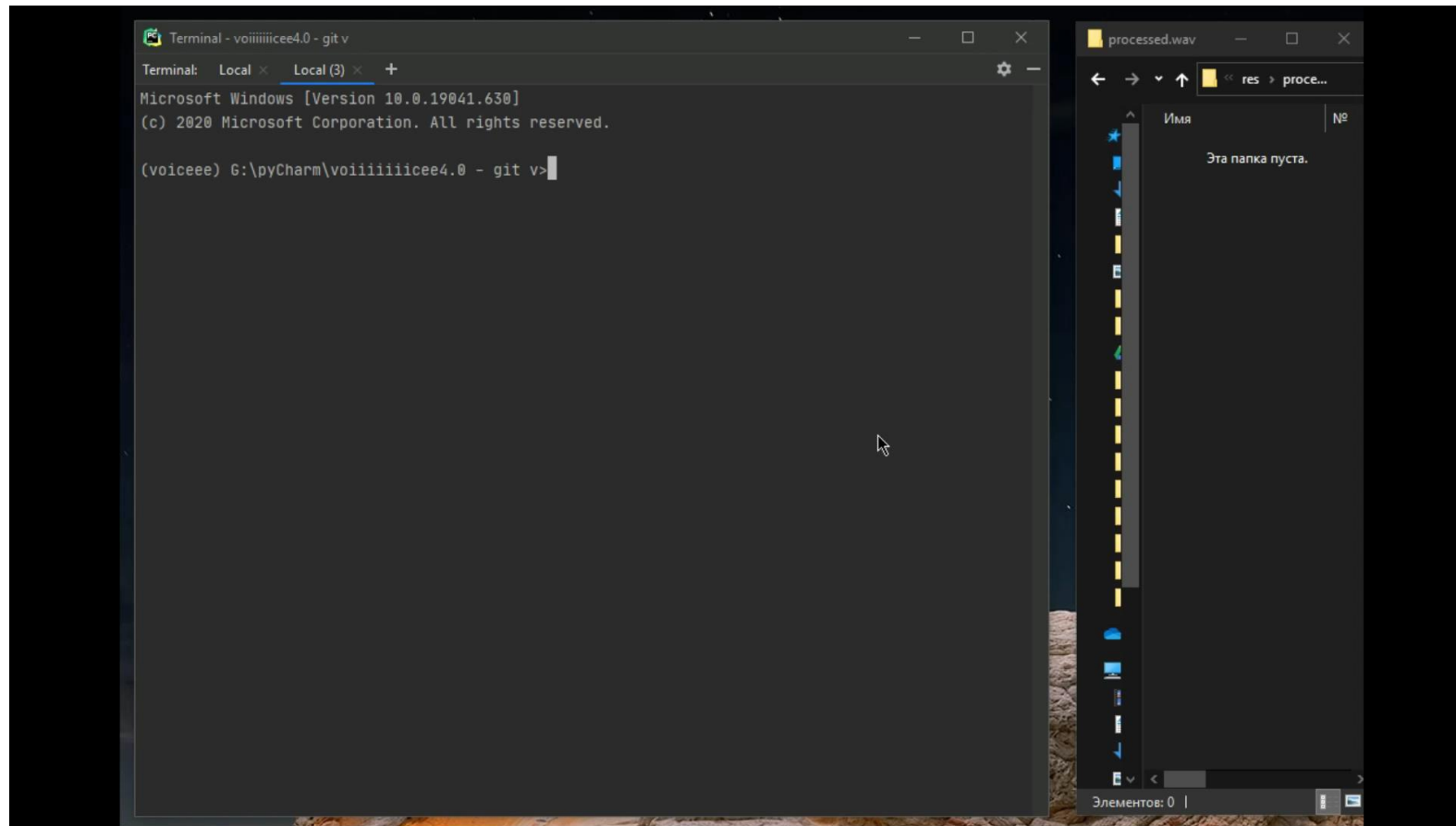


The image shows a screenshot of a Windows terminal window. The title bar reads "Terminal - voiiiiiiiicee4.0 - git v". The terminal has several tabs open, with "Local (6)" selected. The output of the command is as follows:

```
Terminal: Local x Local (3) x Local (2) x Local (4) x Local (5) x Local (6) x +
Microsoft Windows [Version 10.0.19041.630]
(c) 2020 Microsoft Corporation. All rights reserved.

(voiceee) G:\pyCharm\voiiiiiiiicee4.0 - git v>
```

# Достигнутый результат. Демонстрация 2





# 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