**Аналитический обзор использованных методов**

**Описание задачи:** Используя платформу Huggingface, провести анализ моделей или использовать готовую библиотеку Python для разделения аудиодорожки на отдельные персоны и распознать текст, сказанный каждой персоной на аудио.

**Требования:**

1. Разделить аудиодорожку на несколько сегментов, каждый из которых соответствует отдельной персоне.
2. Использовать код из первой задачи для распознавания текста, сказанного каждой персоной.

**Ожидаемые результаты**:

1. Код для разделения аудиодорожки на отдельные персоны.
2. Программа, которая распознает текст, сказанный каждой персоной на аудио.
3. Анализ и обоснование выбранных методов и моделей.

**Методы, использованные в работе**

Программа состоит из конвертера аудио (converter) и модуля обработки аудио-файла (separate\_voices).

**Converter** – предназначен для преобразования аудиофайлов любых типов в тип wav, использует библиотеку pydub. Pydub – это библиотека, которая предоставляет удобные инструменты для работы с аудиофайлами. С помощью этой библиотеки разработчики могут выполнять широкий спектр операций над звуковыми файлами, таких как конвертация форматов, нарезка, склеивание, изменение громкости, применение эффектов и многое другое.

**Separate\_voices** – модуль для сегментации звуковой дорожки по голосам спикеров и преобразования полученных аудио в текст. Основан на библиотеках pyannote.audio, pyannote\_whisper, pywhispercpp.

[Pyannote.audio](https://pyannote.audio/) - это библиотека для обработки аудио данных, разработанная для задач распознавания речи, аудио-классификации, аудио-сегментации и других задач обработки аудио.

Pyannotewhisper - это расширение для [pyannote.audio](https://pyannote.audio/), которое специализируется на анализе шепота. Он предоставляет инструменты для определения и классификации шепота в аудиофайлах.

Pywhispercpp - это дополнительное расширение для pyannotewhisper, которое реализовано на C++ и предоставляет более быстрые и эффективные алгоритмы для анализа шепота.

Для реализации разделения голосов спикеров использовались модели Whisper от OpenAI. Whisper содержит несколько моделей разных размеров и вычислительной мощности: tiny, base, small, medium, large.

Тестирование моделей осуществлялось с использованием аудиофайла записи аудиокниги «Алиса в стране чудес», длительностью 1 минута и двумя дикторами.

Таблица 1. Сравнение моделей whisper

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Модель | Занимаемое дисковое пространство (МБ) | Использование ОЗУ (МБ) | Время обработки аудиофайла (сек) | Замечания |
| tiny | 75 | 273 | 4.734 | Много сливающихся слов, допущены множественные ошибки распознавания слов, есть синтаксические ошибки, пропущены некоторые слова, отсутствуют кавычки в монологах, неправильное разделение текста. |
| base | 142 | 388 | 8.124 | Некоторые слова сливаются в одно, есть ошибки распознавания слов, пропущены некоторые слова, неправильное разделение текста. |
| small | 466 | 852 | 23.565 | Некоторые слова сливаются в одно, есть синтаксические ошибки, неправильное разделение текста. |
| medium | 1536 | 2151 | 73.504 | Неправильное разделение текста. |
| large | 2970 | 3994 | 136.026 | - |

**Транскрибация аудио в текст результаты.**

**Модель tyny:**

0.00 5.56 SPEAKER\_01 У самого выхода она вспомнила, что оставила золотойключек на столе.

5.56 11.28 SPEAKER\_01 А подбежав опять столику обнаружила, что теперь ей никакдоключений дотянуться.

11.28 14.92 SPEAKER\_01 И главное, его было так хорошо видно сквозь стекло.

14.92 34.00 SPEAKER\_01 Она попробовала влезь на стол по ножке, но ножкибыли тоже стеклянные ужасно скольские, и как Алисане старалась, она вновь и вновь съезжала на пол, и наконецна старавше сильно съезжавшись до изнимажения, битьняшко села прямо на пол и заплакала.

34.00 38.44 SPEAKER\_00 Ну вот, еще чего не хватало, сказала Алиса себе довольнострого.

38.44 40.20 SPEAKER\_00 Слезами, горы не поможешь.

40.20 42.80 SPEAKER\_00 Советую тебе перестать с ее минуту.

42.80 48.48 SPEAKER\_00 Алиса вообще всегда давала себе превосходные советы, хотяслушала всех далеко далеко и не всегда.

48.48 53.20 SPEAKER\_00 Иногда она закатывала себе такие выговоры, что Елена моглаудержать, что отслёс.

53.20 60.00 SPEAKER\_00 А как это раз она помнице даже попробовала выдрать себяза уши за то, что жульничали играй самос собой в каком-то.

**Модель base:**

0.00 5.92 SPEAKER\_01 У самого выхода она вспомнила, что оставила золотойключик на столе.

5.92 11.60 SPEAKER\_01 А подбежав опять к столику обнаружила, что теперьей никак до ключа не дотянуться.

11.60 15.56 SPEAKER\_01 И главное, его было так хорошо видно сквозь стекло.

15.56 21.36 SPEAKER\_01 Она попробовала влезь на стол по ножке, но ножки былитоже стеклянные и ужасно скользкие.

21.36 33.84 SPEAKER\_01 И как Алиса не старалась, она вновь и вновь съезжалана пол, и наконец, на старавшийся, насъезжавшись доизнимажения, бить няшко село прямо на пол и заплакала.

33.84 38.08 SPEAKER\_00 "Ну вот, еще чего не хватало", сказала Алиса себедовольно строго.

38.08 40.12 SPEAKER\_00 "Слезами, горе не поможешь.

40.12 42.72 SPEAKER\_00 Советую тебе перестать, сию минуту".

42.72 48.44 SPEAKER\_00 Алиса вообще всегда давала себе превосходные советы,хотя слушалась их далеко, далеко не всегда.

48.44 59.96 SPEAKER\_00 Иногда она закатывала себе такие выговоры, что Елимогла удержаться от слез, а как-то раз она помнится дажепопробовала выдрать себя за уши за то, что жульничалаиграя сама с собой в кратин.

**Модель small:**

0.00 5.60 SPEAKER\_01 У самого выхода она вспомнила, что оставила золотой ключикна столе.

5.60 11.12 SPEAKER\_01 А подбежав опять к столику обнаружила, что теперь ейникак до ключа не дотянуться.

11.12 15.36 SPEAKER\_01 И главное его было так хорошо видно сквозь стекло.

15.36 18.04 SPEAKER\_01 Она попробовала ловлеть на стол по ножке.

18.04 21.40 SPEAKER\_01 Но ножки были тоже стеклянные и ужасно скользкие.

21.40 33.60 SPEAKER\_01 И как Алиса не старалась, она вновь и вновь съезжалана пол и наконец настаравшись и на съезжавшись до изнеможения- битьняшка села прямо на пол и заплакала.

33.60 38.48 SPEAKER\_00 «Ну вот ещё чего не хватало», сказала Алиса себе довольнострого.

38.48 40.16 SPEAKER\_00 «Слезами, Горю, не поможешь.

40.16 42.76 SPEAKER\_00 Советую тебе перестать в сию минуту».

42.76 48.44 SPEAKER\_00 Алиса вообще всегда давала себе превосходные советы,хотя слушала их далеко-далеко не всегда.

48.44 59.96 SPEAKER\_00 Иногда она закатывала себе такие выговоры, что елемогла удержаться от слёз, а как-то раз она помнитсядаже попробовала выдрать себя за уши за то, что жульничалаиграя сама с собой в кроке.

**Модель medium:**

0.00 5.40 SPEAKER\_01 У самого выхода она вспомнила, что оставила золотой ключик на столе.

5.40 11.08 SPEAKER\_01 А подбежав опять к столику, обнаружила, что теперь ей никак до ключа не дотянуться.

11.08 14.68 SPEAKER\_01 И главное, его было так хорошо видно сквозь стекло.

14.68 21.12 SPEAKER\_01 Она попробовала влезть на стол по ножке, но ножки были тоже стеклянные и ужасно скользкие.

21.12 25.88 SPEAKER\_01 И как Алиса не старалась, она вновь и вновь съезжала на пол.

25.88 33.76 SPEAKER\_01 И наконец, настаравшись и насъезжавшись до изнеможения, бедняжка села прямо на пол и заплакала.

33.76 38.04 SPEAKER\_00 — Ну вот еще чего не хватало, — сказала Алиса себе довольно строго.

38.04 42.32 SPEAKER\_00 — Слезами горе не поможешь. Советую тебе перестать сию минуту.

42.32 48.16 SPEAKER\_00 Алиса вообще всегда давала себе превосходные советы, хотя слушалась их далеко-далеко не всегда.

48.16 52.96 SPEAKER\_00 Иногда она закатывала себе такие выговоры, что еле могла удержаться от слез.

52.96 59.96 SPEAKER\_00 А как-то раз она, помнится, даже попробовала выдрать себя за уши за то, что жульничала, играя сама с собой в кроке.

**Модель large:**

0.00 11.16 SPEAKER\_01 У самого выхода она вспомнила, что оставила золотой ключик на столе,а подбежав опять к столику, обнаружила, что теперь ей никак до ключа не дотянуться.

11.16 14.96 SPEAKER\_01 И главное, его было так хорошо видно сквозь стекло.

14.96 21.16 SPEAKER\_01 Она попробовала влезть на стол по ножке, но ножки были тоже стеклянные и ужасно скользкие.

21.16 33.96 SPEAKER\_01 И как Алиса не старалась, она вновь и вновь съезжала на пол,и, наконец, настаравшись и насъезжавшись до изнеможения, бедняжка села прямо на пол и заплакала.

33.96 38.20 SPEAKER\_00 — Ну вот еще чего не хватало, — сказала Алиса себе довольно строго.

38.20 42.48 SPEAKER\_00 — Слезами горе не поможешь. Советую тебе перестать в сию минуту.

42.48 48.24 SPEAKER\_00 Алиса вообще всегда давала себе превосходные советы, хотя слушалась их далеко-далеко не всегда.

48.24 60.00 SPEAKER\_00 Иногда она закатывала себе такие выговоры, что еле могла удержаться от слез,а как-то раз она помнится, даже попробовала выдрать себя за уши за то, что жульничала, играя сама с собой в кракет.