**Практическая работа №2**

**Потоковая аналитика эмоциональных оттенков речи человека**

**Цель работы:**

Выполнение практической работы направлено на изучение современных программных средств определения эмоционального фона речи человека в голосе и в тексте.

**Порядок работы:**

1. Определение эмоционального фона речи человека в текстовой форме с использованием модуля *Text* библиотеки *Aniemore* для *Python*. Проверка модели *rubert-tiny-emotion-russian-cedr-m7*. Программа определила эмоциональный фон как «нейтральный» с точностью 38%, как показано на рисунке 1.



Рисунок 1 - Проверка модели архитектуры Bert\_Tiny

2. Проверка модели *rubert-tiny2-russian-emotion-detection*. Программа определила эмоциональный фон как «гневный» с точностью 80%, как показано на рисунке 2.

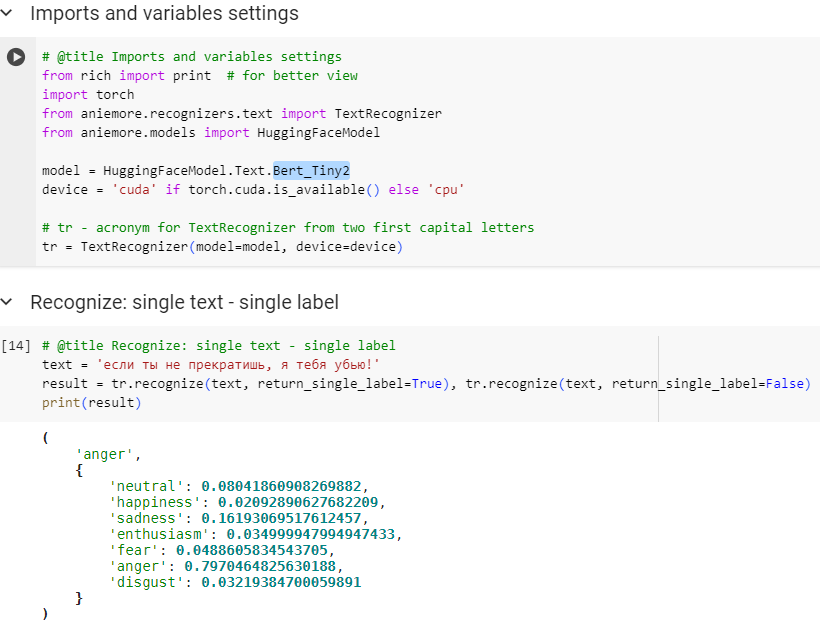
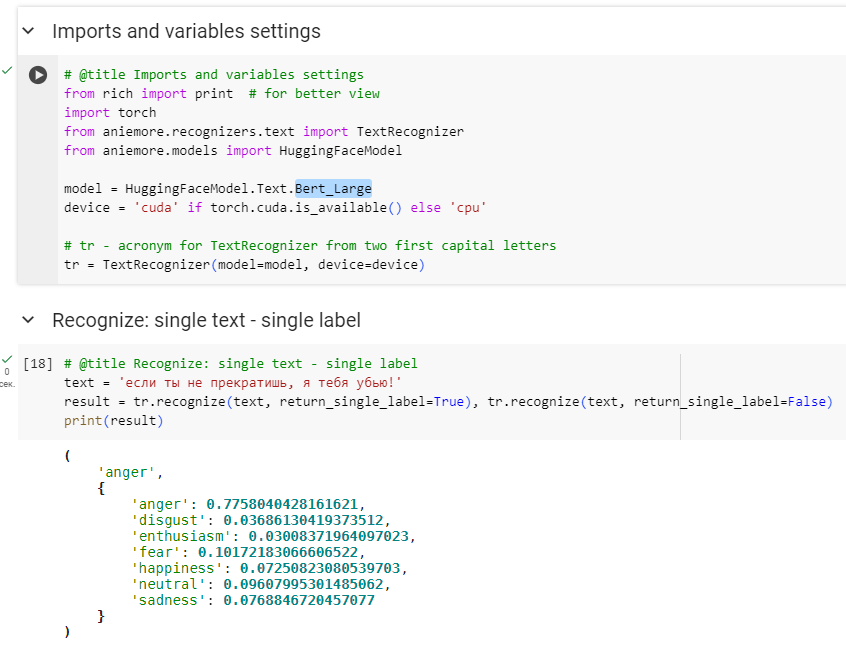
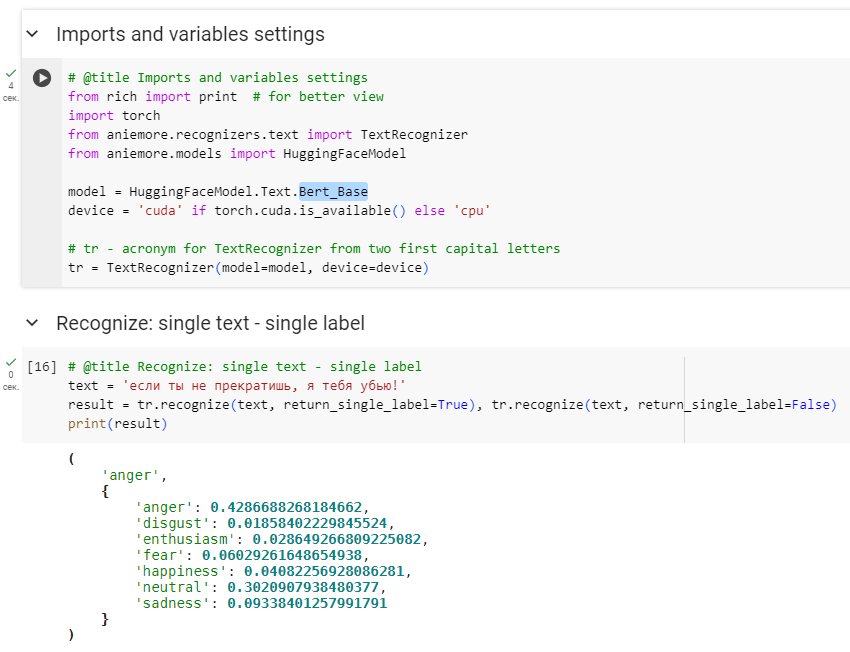


Рисунок 2 - Проверка модели архитектуры Bert\_Tiny2

3. Проверка модели *rubert-large-emotion-russian-cedr-m7*. Программа определила эмоциональный фон как «гневный» с точностью 78%, как показано на рисунке 3.

Рисунок 3 - Проверка модели архитектуры Bert\_Large

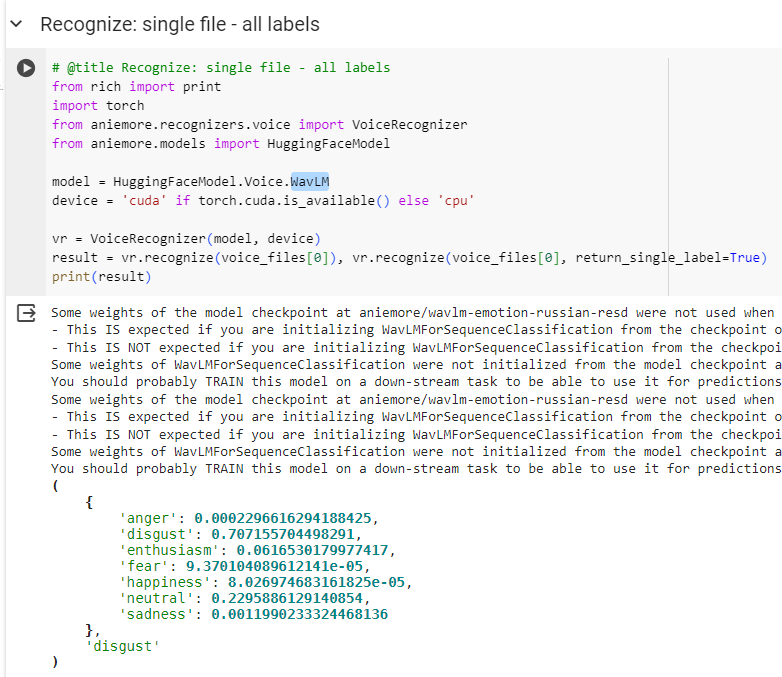
4. Проверка модели *rubert-base-emotion-russian-cedr-m7*. Программа определила эмоциональный фон как «гневный» с точностью 43%, как показано на рисунке 4.

Рисунок 4 - Проверка модели архитектуры Bert\_Base

5. Подготовка к определению эмоционального фона речи человека в голосовой форме с использованием модуля *Voice* библиотеки *Aniemore* для *Python*. Загрузка файла голосовой записи представлена на рисунке 5.

Рисунок 5 - Загрузка звукового файла

6. Проверка модели *wavlm-emotion-russian-resd*. Программа определила эмоциональный фон как «отвращенный» с точностью 71%, как показано на рисунке 6.

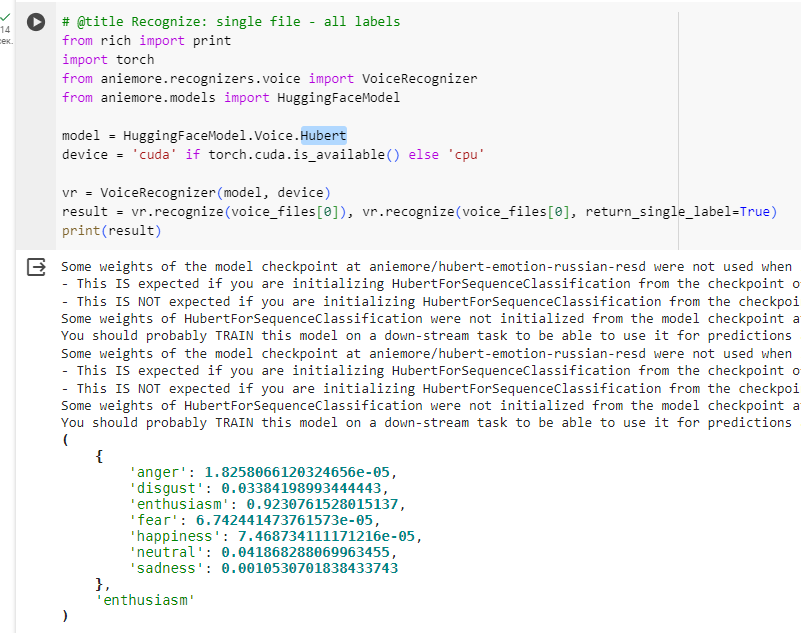
Рисунок 6 - Проверка модели архитектуры WavLM

7. Проверка модели *wav2vec2-xlsr-53-russian-emotion-recognition*. Программа определила эмоциональный фон как «нейтральный» с точностью 99%, как показано на рисунке 7.



Рисунок 7 - Проверка модели архитектуры Wav2Vec2

8. Проверка модели *hubert-emotion-russian-resd*. Программа определила эмоциональный фон как «воодушевленный» с точностью 92%, как показано на рисунке 8.

Рисунок 8 - Проверка модели архитектуры Hubert

9. Проверка модели *unispeech-sat-emotion-russian-resd*. Программа определила эмоциональный фон как «воодушевленный» с точностью 64%, как показано на рисунке 9.

Рисунок 9 - Проверка модели архитектуры UniSpeech

10. Результаты исследований приведены в таблице 1.

Таблица 1 - Сравнение точности

|  |  |  |
| --- | --- | --- |
| Модель | Заявленная точность | Подтвержденная точность |
| Голосовые модели |  |  |
| [**wav2vec2-xlsr-53-russian-emotion-recognition**](https://huggingface.co/Aniemore/wav2vec2-xlsr-53-russian-emotion-recognition) | 73% | 99% |
| [**wav2vec2-emotion-russian-resd**](https://huggingface.co/Aniemore/wav2vec2-emotion-russian-resd) | 75% | 99% |
| [**wavlm-emotion-russian-resd**](https://huggingface.co/Aniemore/wavlm-emotion-russian-resd) | 82% | 71% |
| [**hubert-emotion-russian-resd**](https://huggingface.co/Aniemore/hubert-emotion-russian-resd) | 75% | 92% |
| [**unispeech-sat-emotion-russian-resd Copied**](https://huggingface.co/Aniemore/unispeech-sat-emotion-russian-resd) | 72% | 64% |
| [**wavlm-bert-base**](https://huggingface.co/Aniemore/wavlm-bert-base-s-emotion-russian-resd) | 81% | 71% |
| [**wavlm-bert-fusion**](https://huggingface.co/Aniemore/wavlm-bert-fusion-s-emotion-russian-resd) | 83% | 71% |
| Текстовые модели |  |  |
| [**rubert-base-emotion-russian-cedr-m7**](https://huggingface.co/Aniemore/rubert-base-emotion-russian-cedr-m7) | 74% | 43% |
| [**rubert-tiny2-russian-emotion-detection**](https://huggingface.co/Aniemore/rubert-tiny2-russian-emotion-detection) | 85% | 80% |
| [**rubert-large-emotion-russian-cedr-m7**](https://huggingface.co/Aniemore/rubert-large-emotion-russian-cedr-m7) | 76% | 78% |
| [**rubert-tiny-emotion-russian-cedr-m7**](https://huggingface.co/Aniemore/rubert-tiny-emotion-russian-cedr-m7) | 72% | 38% |

**Вывод:** в ходе работы ознакомились с возможностями современных программных средств определения эмоционального фона речи человека в голосе и в тексте, получили навыки работы с открытой библиотекой искусственного интеллекта *Aniemore*, в результате проверки моделей библиотеки определили, что в модуле *Text* наивысшую точность показали модели архитектуры *Bert\_Tiny2* (80%) и *Bert\_Large* (78%), в модуле *Voice* — *Wav2Vec2* (99%) и *Hubert* (92%).