



Russian dataset description

Tags	Dataset Emo Old
Date	@22/08/2022
Status	Done

Path: /media/storage/data/raw/emo_audio/russian_emo_dataset_for_labelling

Real samples with id from 0 to 1115

Re-labelling validation samples with id from 1116 to 1301

Metadata:

<https://s3-us-west-2.amazonaws.com/secure.notion-static.com/24ebe22e-e9a6-49e6-85c3-17e69467060f/metadata.json>

Relabelled path:

/media/storage/data/raw/emo_audio/upwork_emo_relabeling_2022-08-19_russian_dataset

Relabelling validation samples info

Taken from PP dataset with threshold 4 test set (path:

/media/storage_2/data/processed/emo_audio/emo_audio_pp_relabelled_threshold_4_4_labelers_no_other_with_wavs)

All samples from:

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/2c2a868a-5c4e-4d18-b808-6b52287a1835/test_dataset.json

Relabelled data quality (validation samples accuracy) by labeller

Iryna: 81.2%

Rime: 90.3%

Gabriela: 96.8% (previously: 54.8%)

Processed data

Path on asya-5:

/media/storage_2/data/processed/emo_audio/emo_audio_ru_relabelled_threshold_2

Sample counts (both test and train):

anger samples: 146

happiness samples: 125

other samples: 1129

sadness samples: 130

In test set each class has 22 samples