

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра математического обеспечения и применения ЭВМ

ОТЧЕТ
по лабораторной работе №2
по дисциплине «Анализ звука и голоса»
ТЕМА: КЛАССИФИКАЦИЯ РЕЧЕВЫХ СИГНАЛОВ ПО ЭМОЦИЯМ

Студентка гр. 6304

Преподаватель

Вероха В. Н.

Рыбин С. В.

Санкт-Петербург

2021

Цель работы.

Классификация данных речевых сигналов по эмоциям.

Описание данных.

В выборке присутствует эмоциональная речь по 5 классам.

- Angry.
- Chilled.
- Happy.
- Neutral.
- Sad.

Выполнение работы.

1. Подключены необходимы библиотеки. Результаты на рис. 1.

```
# Подключение библиотек
import librosa
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import os
from PIL import Image
import pathlib
import csv

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler

import keras
from keras import models
from keras import layers

from tqdm.notebook import tqdm

import warnings
warnings.filterwarnings('ignore')
```

Рисунок 1 — Подключение библиотек

2. Извлечены функции из аудио с использованием Librosa, а также записаны в датафрейм. Фрагмент полученного датафрейма представлен на рис. 2.

	filename	chroma_stft	rmse	spectral_centroid	spectral_bandwidth	rolloff	zero_crossing_rate	mfcc1	mfcc2	mfcc3	...	mfcc12	
0	001.wav	0.449320	0.061361	3079.059976	2693.275778	6049.869999	0.229912	-343.794189	30.688906	-15.840018	...	-2.452909	-4
1	002.wav	0.506070	0.057922	2367.790483	2384.544363	4229.604727	0.131504	-366.752838	28.327007	-18.157299	...	-5.445252	-0
2	003.wav	0.432045	0.023526	2591.701724	2619.853698	5499.470355	0.163465	-474.658630	34.974876	-7.841065	...	-2.695741	-3
3	004.wav	0.482492	0.058055	2525.448793	2413.145172	4952.198543	0.136912	-349.897095	43.330460	-18.259544	...	0.638473	-2
4	005.wav	0.388032	0.058289	3115.824434	2492.394799	5793.514316	0.182514	-388.422485	34.102715	-20.510458	...	-5.720950	-7

Рисунок 2 — Фрагмент извлеченных функций

3. Извлечены типы эмоциональной окраски из названий папок и дописаны отдельным столбцом “label” к датафрейму. Результаты представлены на рис. 3.

zero_crossing_rate	mfcc1	mfcc2	mfcc3	...	mfcc12	mfcc13	mfcc14	mfcc15	mfcc16	mfcc17	mfcc18	mfcc19	mfcc20	label
0.229912	-343.794189	30.688906	-15.840018	...	-2.452909	-4.215964	-0.071347	-3.416410	-5.418313	-2.700627	1.247998	0.751982	4.828207	angry
0.131504	-366.752838	28.327007	-18.157299	...	-5.445252	-0.062126	0.066857	-6.336366	1.053317	-7.168151	-1.407118	1.191806	2.260072	angry
0.163465	-474.658630	34.974876	-7.841065	...	-2.695741	-3.054490	-1.559706	-3.396726	-1.685482	-0.118342	1.202329	1.822596	7.559172	angry
0.136912	-349.897095	43.330460	-18.259544	...	0.638473	-2.003618	2.220423	-5.797532	3.321220	-0.829450	2.314820	5.491830	8.821283	angry
0.182514	-388.422485	34.102715	-20.510458	...	-5.720950	-7.264971	-5.371935	-2.434712	-7.158056	-0.002204	2.571340	-0.628871	3.901053	angry

Рисунок 3 — Типы акустических шумов

4. Полученные данные извлеченных функций преобразованы для обучения. Результаты представлены на рис. 4.

```
array([[ -0.40763205,  3.80280416,  0.66277582, ...,  0.64427888,
         0.90513809,  1.59496332],
       [  0.4640073 ,  3.54989011, -1.36836336, ..., -0.22661084,
         1.04311311,  0.8381416 ],
       [ -0.6729604 ,  1.02057738, -0.72895045, ...,  0.62929918,
         1.24099467,  2.39977089],
       ...,
       [  1.00352492, -0.56213561,  0.57679699, ..., -0.54916428,
         0.05821169,  0.23895281],
       [  1.07467689, -0.55117232,  0.79667249, ..., -1.17219894,
        -0.39330505,  0.01063524],
       [ -2.3324741 , -0.50044242, -2.2444935 , ..., -0.39164112,
        -1.10656368, -0.48902932]])
```

Рисунок 4 — Вектор X для обучения

	filename	label	predict_score	predict_classes
0	001.wav	angry	0.999994	angry
1	002.wav	angry	0.999511	angry
2	003.wav	angry	0.999863	angry
3	004.wav	angry	0.999999	angry
4	005.wav	angry	0.998273	angry

Рисунок 7 — Фрагмент файла с результатами

Выводы.

В результате проделанной лабораторной работы были получены навыки программирования на языке Python. Изучена задача классификации речевых сигналов по эмоциям.

Создана модель с точностью обучения — 89.37%. А также составлен файл с результатами.