

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

**ОТЧЕТ**  
**по лабораторной работе №1**  
**по дисциплине «Анализ звука и голоса»**  
**ТЕМА: КЛАССИФИКАЦИЯ АКУСТИЧЕСКИХ ШУМОВ**

Студентка гр. 6304

Преподаватель

\_\_\_\_\_  
\_\_\_\_\_

Вероха В.Н.

Рыбин С.В.

Санкт-Петербург

2021

## Цель работы.

Классификация данных акустических событий.

## Описание данных.

В выборке присутствует 8 типов акустических шумов:

- background (background\_....);
- bags (bags\_..., bg\_..., t\_bags\_...);
- door (door\_..., d\_..., t\_door\_...);
- keyboard (keyboard\_..., t\_keyboard\_..., k\_...);
- knocking\_door (knocking\_door\_..., tt\_kd\_...);
- ring (ring\_..., t\_ring\_...);
- speech;
- tool.

## 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
0	background_0001.wav	0.624490	0.000158	3034.544804	2307.485058	5953.745943	0.250282	-738.071716	93.217445	-33.399258
1	background_0002.wav	0.745098	0.000071	3186.990847	2310.177615	6053.988874	0.272419	-789.601990	75.197586	-31.433498
2	background_0003.wav	0.682624	0.000073	3389.104803	2308.878995	6213.184290	0.304189	-778.028931	73.797752	-33.068264
3	background_0004.wav	0.760644	0.000089	3199.924891	2311.821234	6079.331513	0.279971	-780.314270	65.745621	-22.033247
4	background_0005.wav	0.772042	0.000051	3215.602229	2310.978107	6092.315856	0.278374	-797.359863	72.981621	-32.363747

Рисунок 2 — Фрагмент извлеченных функций

3. Извлечены типы акустических шумов из названий файлов 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
0.250282	-738.071716	93.217445	-33.399258	...	-6.226457	12.462728	-12.907367	12.445766	-9.406123	10.940209	-3.830426	3.890665	-1.316001	background
0.272419	-789.601990	75.197586	-31.433498	...	-5.241376	10.665572	-9.677876	11.937750	-8.288611	9.191567	-3.542699	4.432508	0.822094	background
0.304189	-778.028931	73.797752	-33.068264	...	-7.054873	9.454710	-12.917912	11.743085	-9.146166	10.361667	-3.550290	4.598960	0.036682	background
0.279971	-780.314270	65.745621	-22.033247	...	-4.947610	8.886817	-7.817257	9.879170	-6.393075	6.643980	-2.605803	2.302400	1.334502	background
0.278374	-797.359863	72.981621	-32.363747	...	-4.422258	10.689495	-9.336390	11.430396	-8.628820	8.254793	-4.460248	3.480893	-0.209078	background

Рисунок 3 — Типы акустических шумов

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

```
array([[ 0.90503261, -0.94819926,  0.53226763, ...,  0.63640432,
        0.07001937,  0.43826544],
       [ 1.82503306, -0.94955348,  0.71722489, ...,  0.68123133,
        0.15589143,  0.83070476],
       [ 1.3484812 , -0.94951433,  0.96244242, ...,  0.6800486 ,
        0.18227087,  0.68654525],
       ...,
       [ 0.14468148, -0.8897629 , -1.17219616, ...,  0.62378556,
       -1.3317461 , -0.17756825],
       [-0.00623053, -0.86132836, -0.90685575, ...,  0.35638808,
       -1.22963515, -0.34344363],
       [-0.05534983, -0.87700982, -1.04255573, ...,  0.38418734,
       -1.3485414 , -0.29691326]])
```

Рисунок 4 — Вектор X для обучения

5. Составлен вектор данных с ярлыками возможных типов шумов. Результаты представлены на рис. 5.

```
y
array([0, 0, 0, ..., 5, 5, 5])
```

Рисунок 5 — Вектор y для обучения

6. Данные для моделирования разделены на выборки для обучения и тестов в соотношении 80/20.

7. Создана модель, архитектура которой представлена на рис. 6.

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
dense_5 (Dense)	(None, 512)	13824
dense_6 (Dense)	(None, 256)	131328
dense_7 (Dense)	(None, 128)	32896
dense_8 (Dense)	(None, 64)	8256
dense_9 (Dense)	(None, 10)	650

=====  
Total params: 186,954  
Trainable params: 186,954  
Non-trainable params: 0

Рисунок 6 — Архитектура модели

8. Точность построенной модели составила 96.03%.

9. Сформирован файл с результатами тестовой выборки. Фрагмент файла представлен на рис. 7.

Out[69]:

	filename	predict_score	predict_classes
0	01.wav	1.495658e-09	tool
1	02.wav	1.077310e-07	tool
2	03.wav	1.838001e-07	tool
3	04.wav	2.205760e-08	tool
4	05.wav	4.044625e-08	keyboard
...	...	...	...
427	95.wav	7.251053e-16	bags
428	96.wav	2.201824e-07	tool
429	97.wav	1.065607e-08	tool
430	98.wav	2.204313e-04	bags
431	99.wav	4.431747e-05	keyboard

432 rows x 3 columns

Рисунок 7 — Фрагмент файла с результатами

## **Выводы.**

В результате проделанной лабораторной работы были получены навыки программирования на языке Python. Изучена задача классификации акустических шумов.

Создана модель с точностью обучения — 99.83%, тестирования — 95.75%. А также составлен файл с результатами.