

ПРОЕКТ

Классификация таймфреймов аудио по типу музыкального инструмента



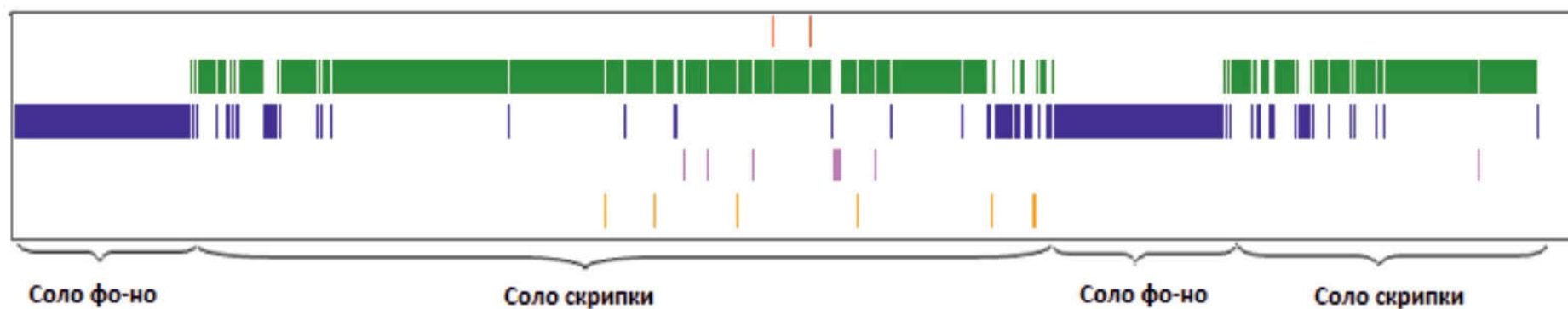
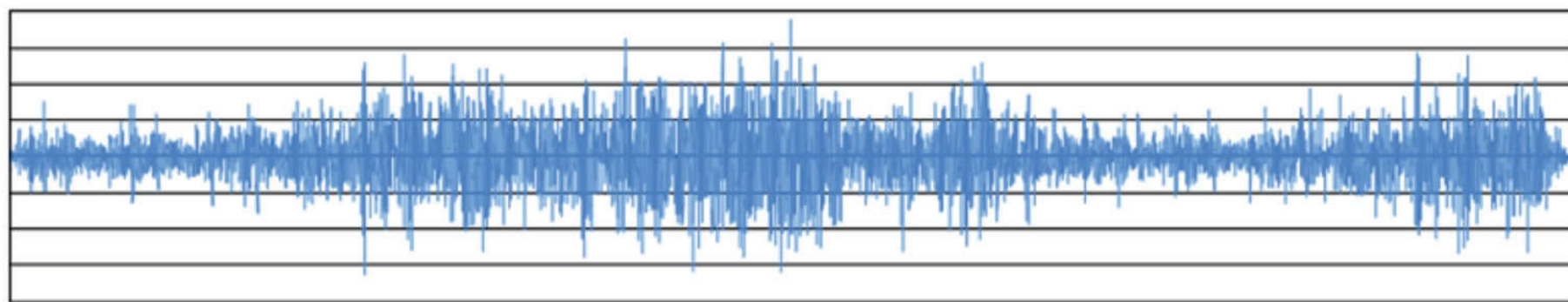
Витомсков А.В. КЭ-220



Введение

- Задача распознавания музыкальных инструментов востребована при обработке сложных музыкальных сигналов.
- Данная задача относится к классу задач Music Information Retrieval (MIR).
- Результаты работы могут быть использованы для аннотирования медиаконтента, сегментации аудиосигналов и решения других задач обработки звука.





University of Iowa Electronic Music Studios

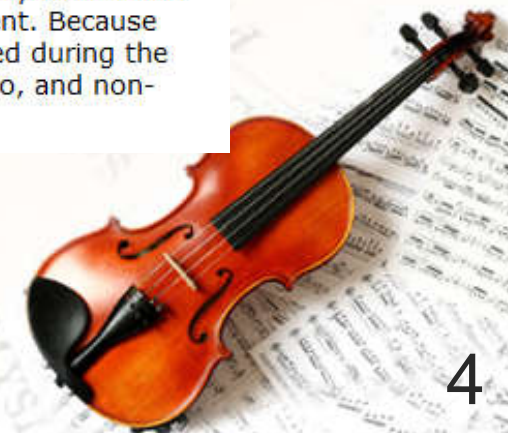
[Home](#) [Musical Instrument Samples](#) [Facilities/Archives](#) [Sounds/Images](#) [Events](#) [Teaching](#) [People](#) [Giving](#)



Musical Instrument Samples

The University of Iowa Musical Instrument Samples (MIS) are created by Lawrence Fritts, Director of the Electronic Music Studios and Professor of Composition at the University of Iowa. Since 1997, these recordings have been freely available on this website and may be downloaded and used for any projects, without restrictions. These are used by musicians, application developers, teachers, students, and researchers. These have been used in over 270 published research articles and books. Please let me know if these have been helpful in your research and music projects.

Until 2011, the recordings were made with single Neumann KM 84 cardioid condenser microphone at 16/44.1. They were edited into chromatic scales played note-by-note at pp, mf, and ff dynamic levels throughout the range of the instrument. Because the recordings represent the complete dynamic structure of the instrument, input volume levels were not changed during the recording session. Some instruments were played with more than one technique, including arco, pizzicato, vibrato, and non-vibrato.



Выбор датасета

- Использована база данных музыкальных инструментов Университета Айовы <http://theremin.music.uiowa.edu>
- База данных образцов включает струнные инструменты, духовые инструменты, медные духовые инструменты, фортепиано
- Полный диапазон звучания
- Инструменты записаны в безэховой камере



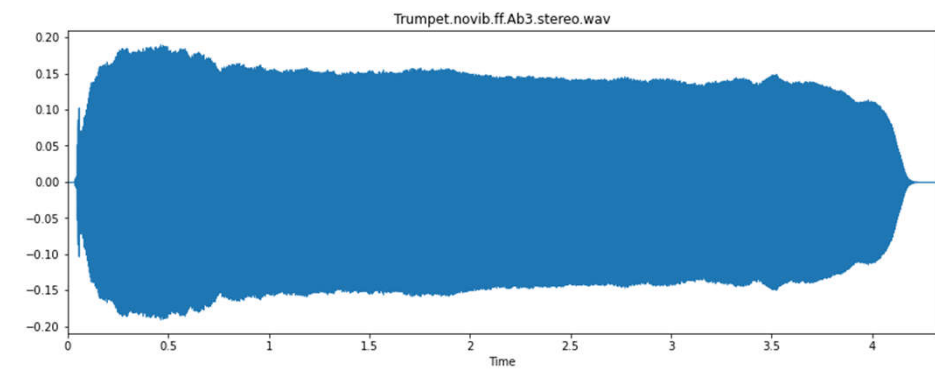
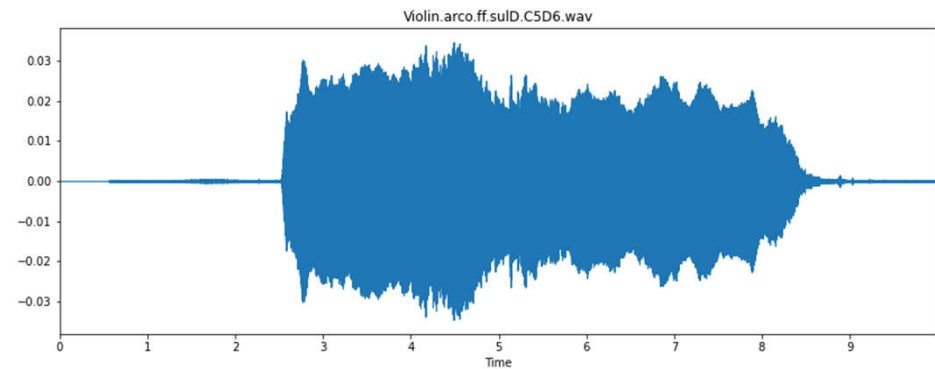
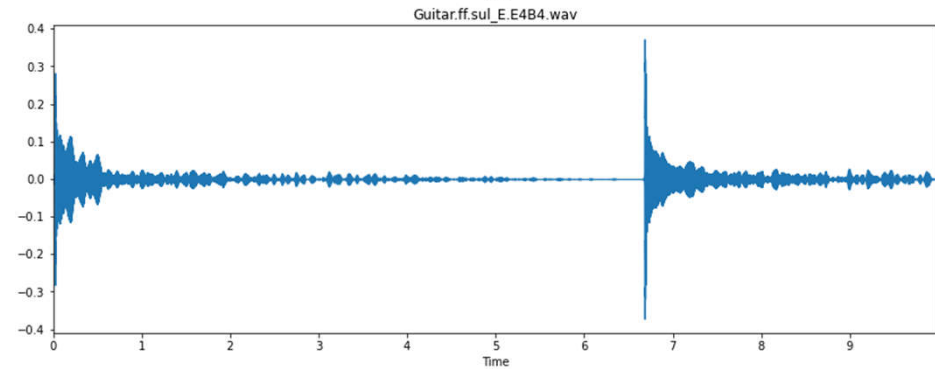
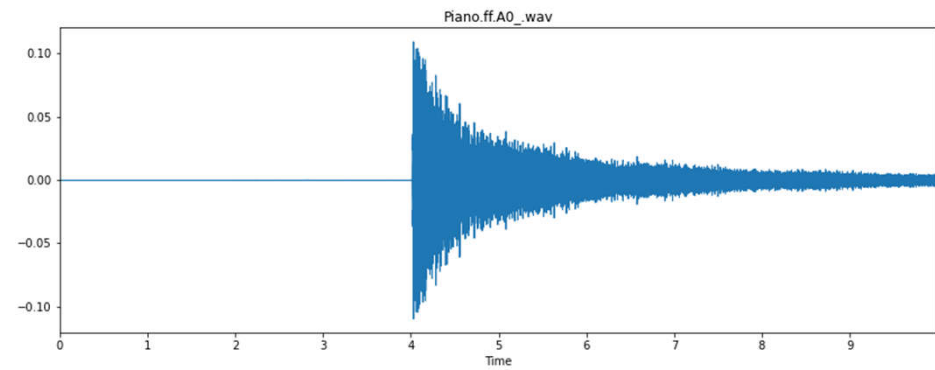
ДО ОБРАБОТКИ

Фортепиано

Гитара

Скрипка

Труба



Предобработка звуковых файлов

Битовая глубина сэмпла: 8 бит,
Частота дискретизации: 22050 Гц,
Каналов: 1 (моно)

```
#вырезать тишину
chunks = split_on_silence(sound,
# Specify that a silent chunk must be at least 2 seconds or 2000 ms long.
min_silence_len = 200,
# Consider a chunk silent if it's quieter than -16 dBFS.
silence_thresh = -37,
# keep 200 ms of leading/trailing silence
keep_silence = 0)

if chunks != []:
    arr_of_chunks = chunks[0]
    for chunk in chunks[1:]:
        arr_of_chunks = arr_of_chunks + chunk
    sound = arr_of_chunks

sound = sound.set_channels(1)
sound = sound.set_sample_width(1)
sound = sound.set_frame_rate(sr)

filepath_wav = os.path.join(new_root, Path(file).stem + '.wav')
sound.export(filepath_wav, format="wav")
return
```



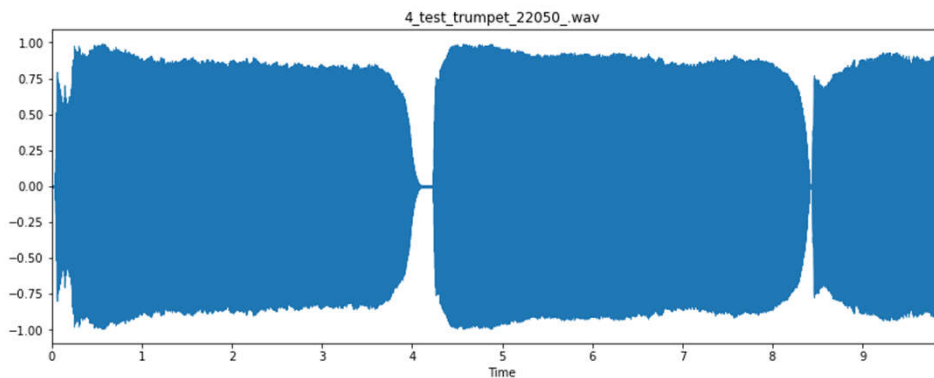
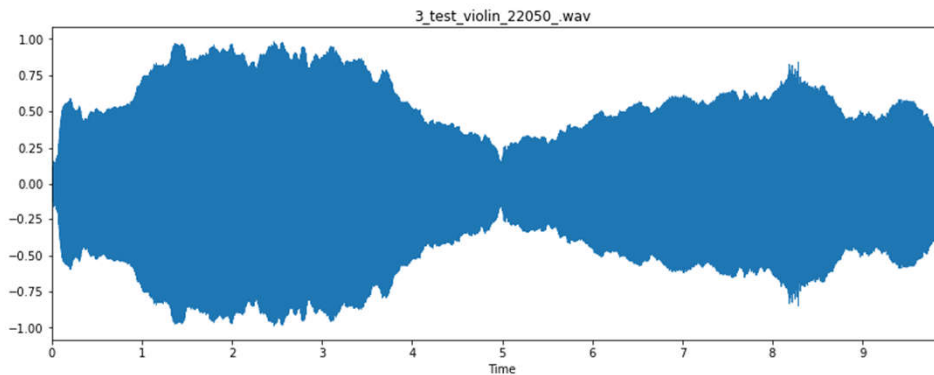
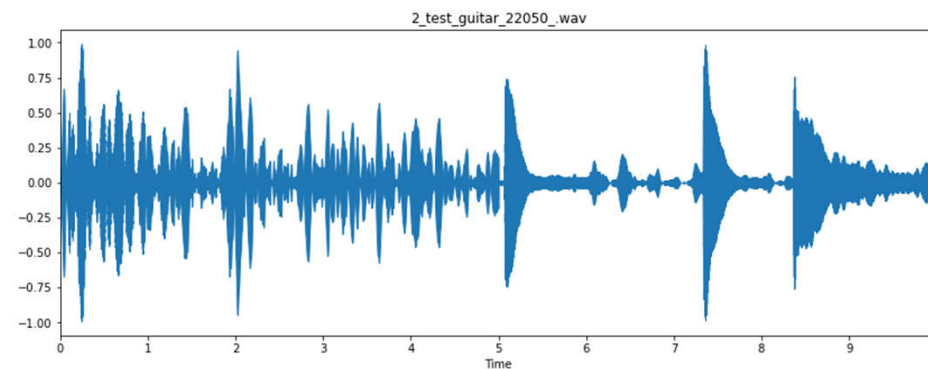
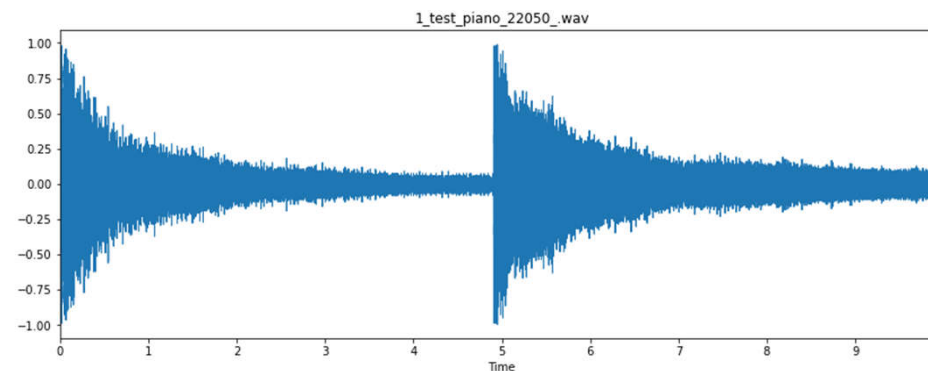
ПОСЛЕ ОБРАБОТКИ

Фортепиано

Гитара

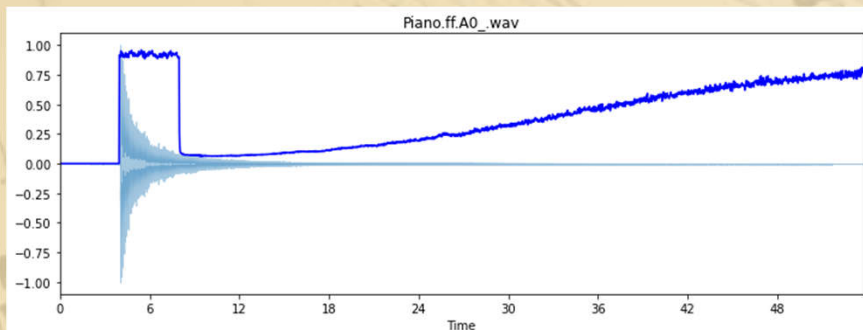
Скрипка

Труба

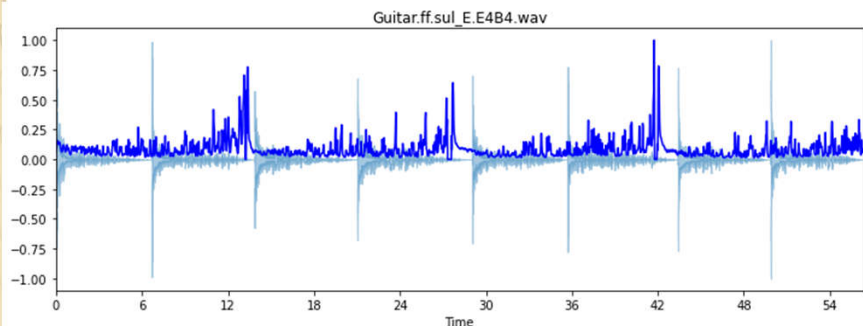


Какие признаки выбрать?

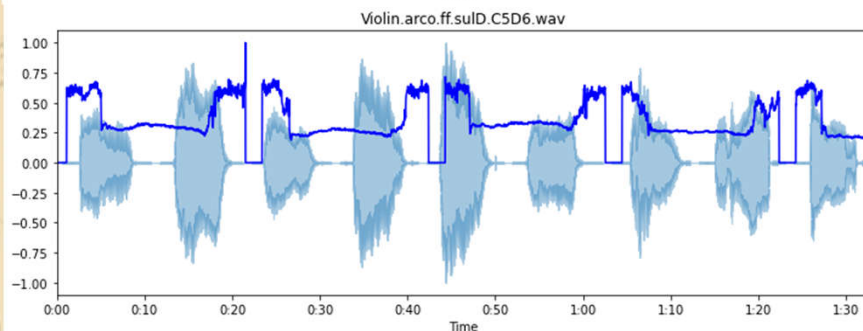
СПЕКТРАЛЬНЫЙ ЦЕНТРОИД
Фортепиано



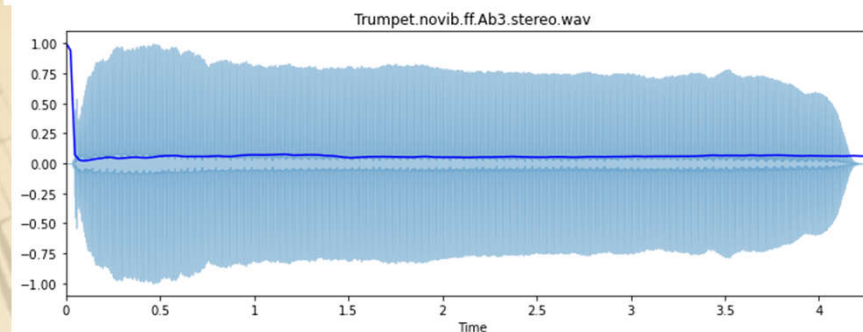
Гитара



Скрипка

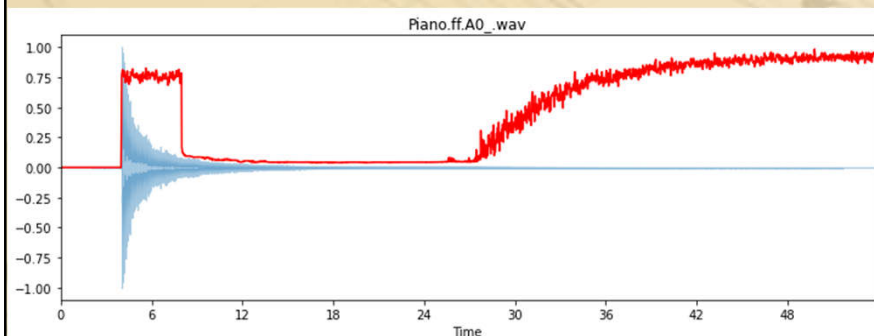


Труба

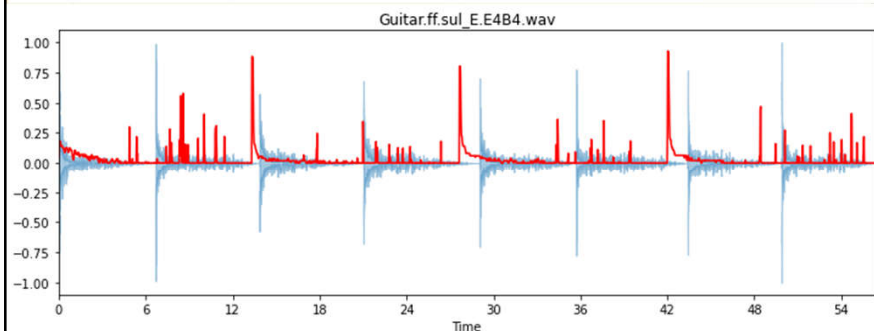


Какие признаки выбрать?

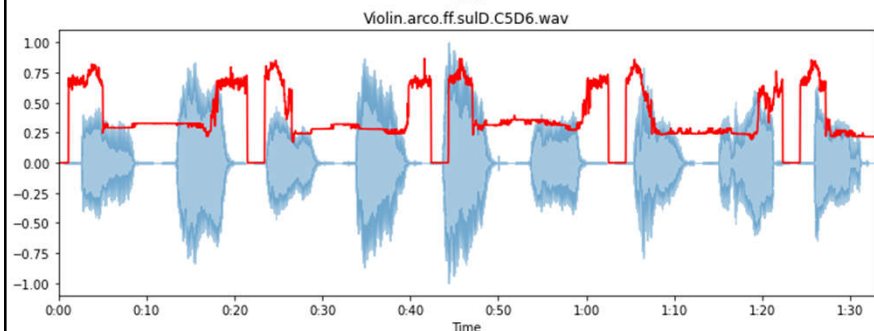
СПЕКТРАЛЬНЫЙ СПАД
Фортепиано



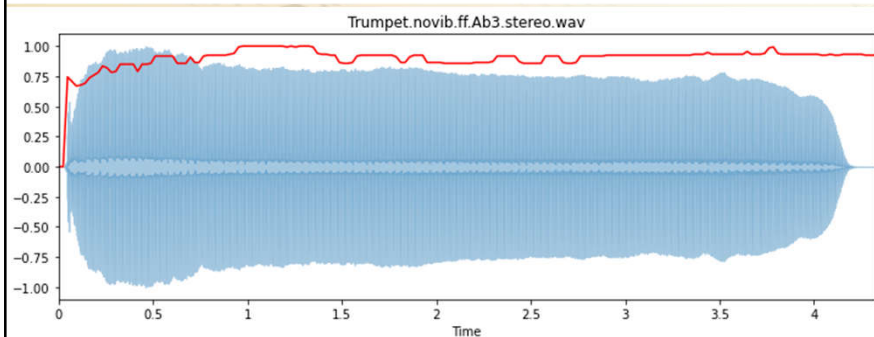
Гитара



Скрипка



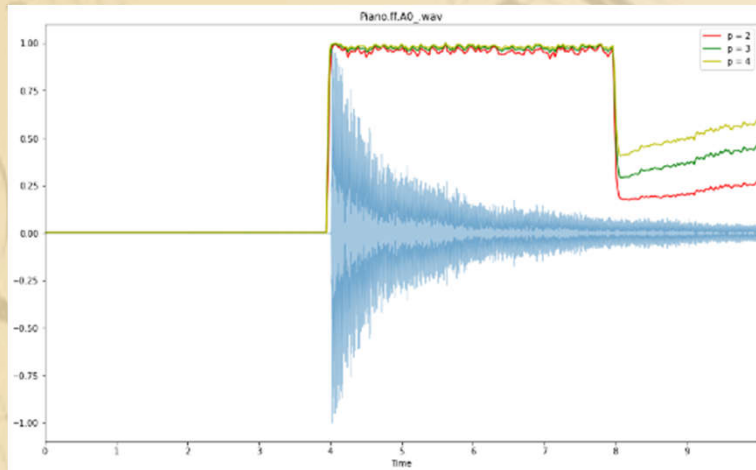
Труба



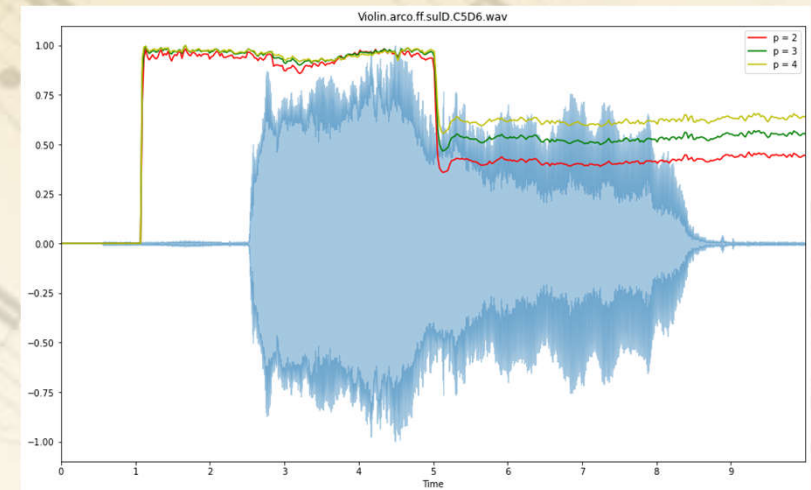
Какие признаки выбрать?

СПЕКТРАЛЬНАЯ ШИРИНА

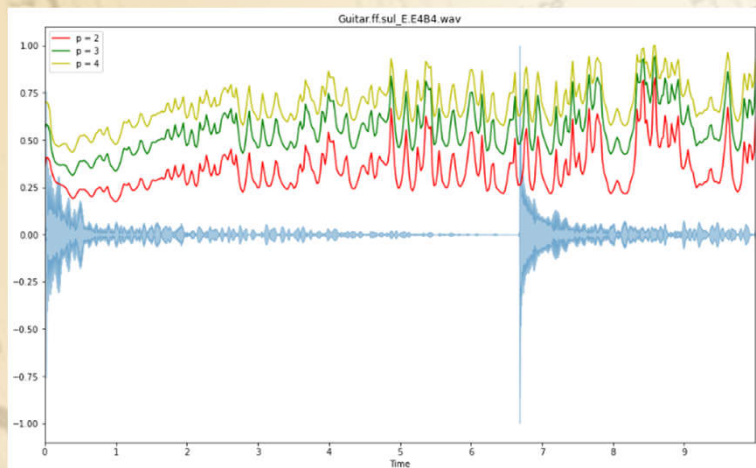
Фортепиано



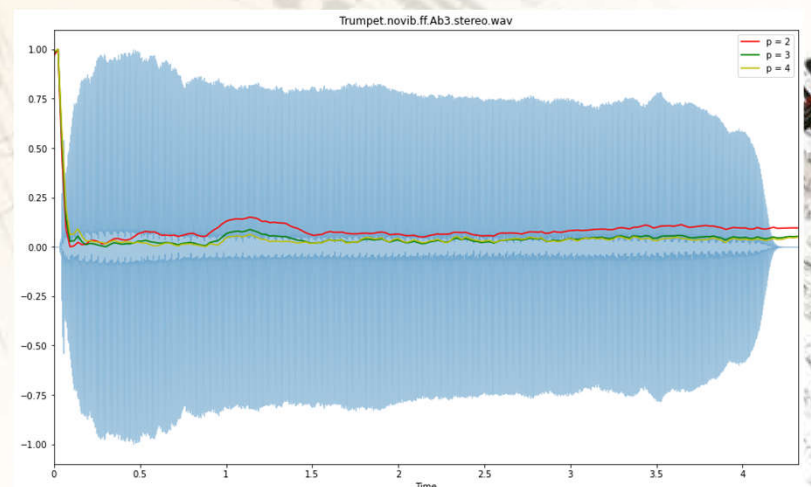
Скрипка



Гитара



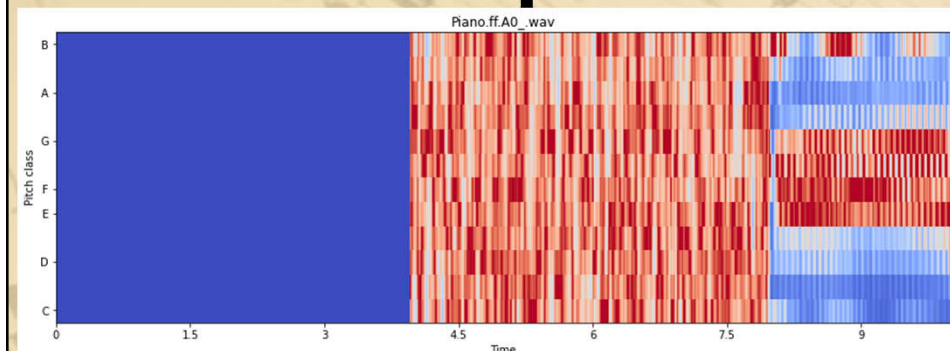
Труба



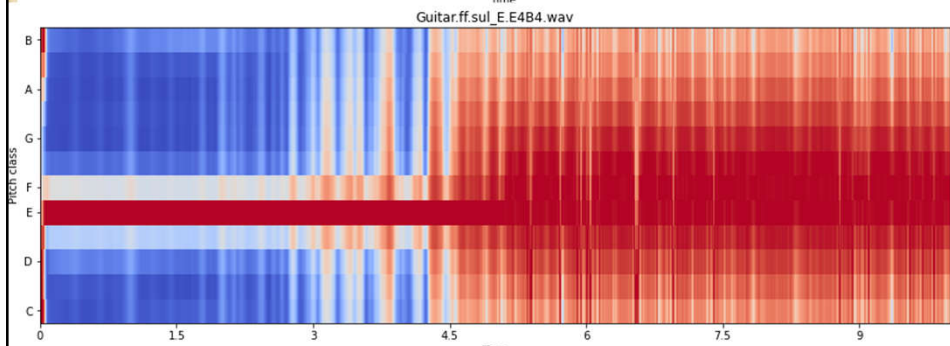
Какие признаки выбрать?

ЦВЕТНОСТЬ

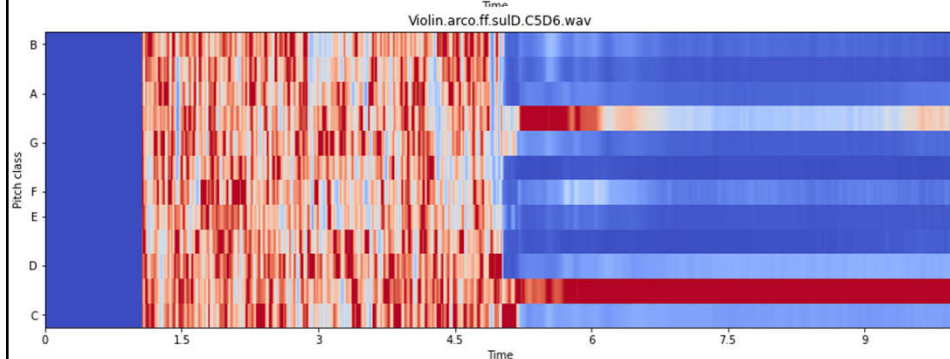
Фортепиано



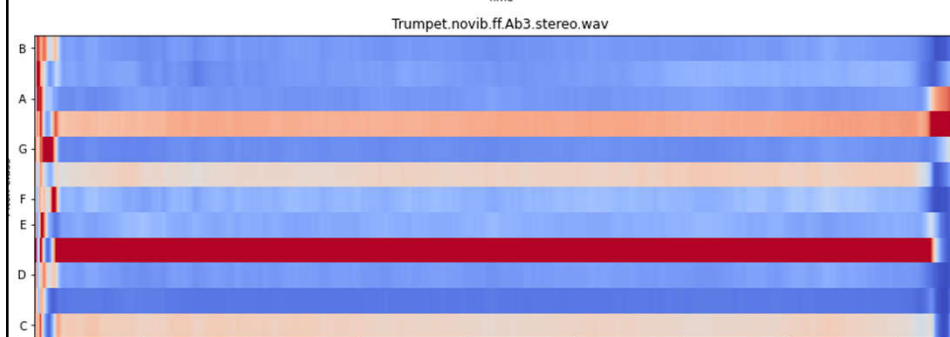
Гитара



Скрипка



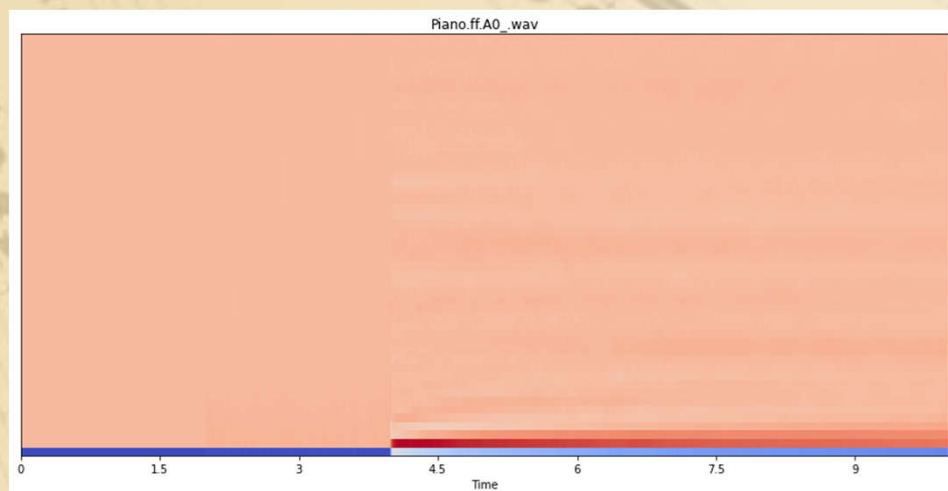
Труба



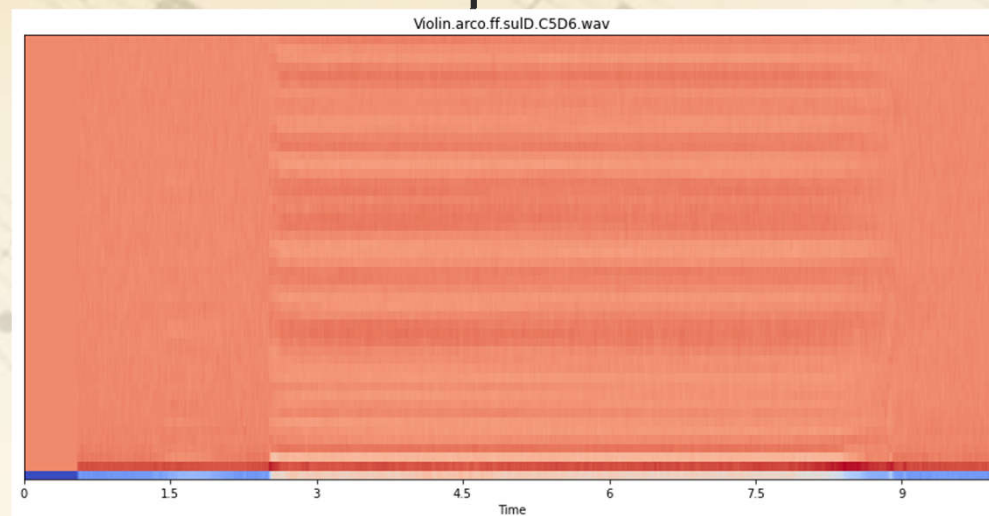
Какие признаки выбрать?

МЕЛ-КЕПСТРАЛЬНЫЕ
КОЭФИЦИЕНТЫ

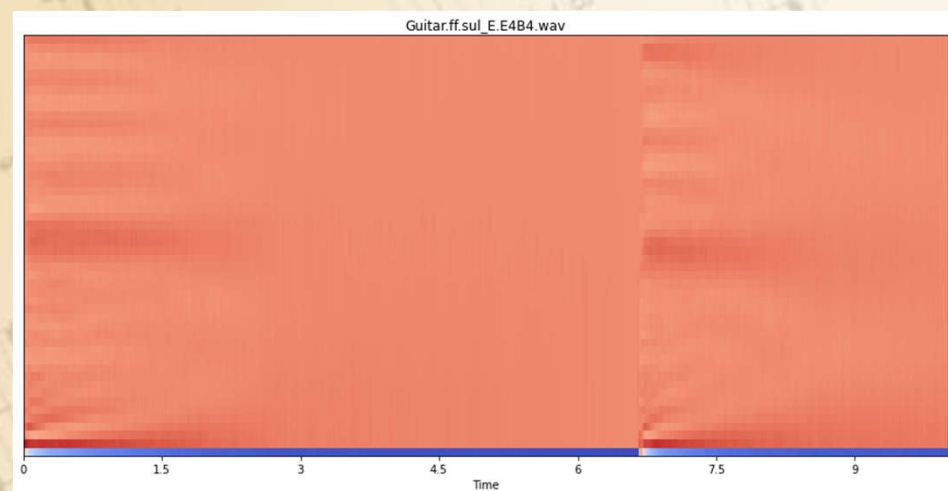
Фортепиано



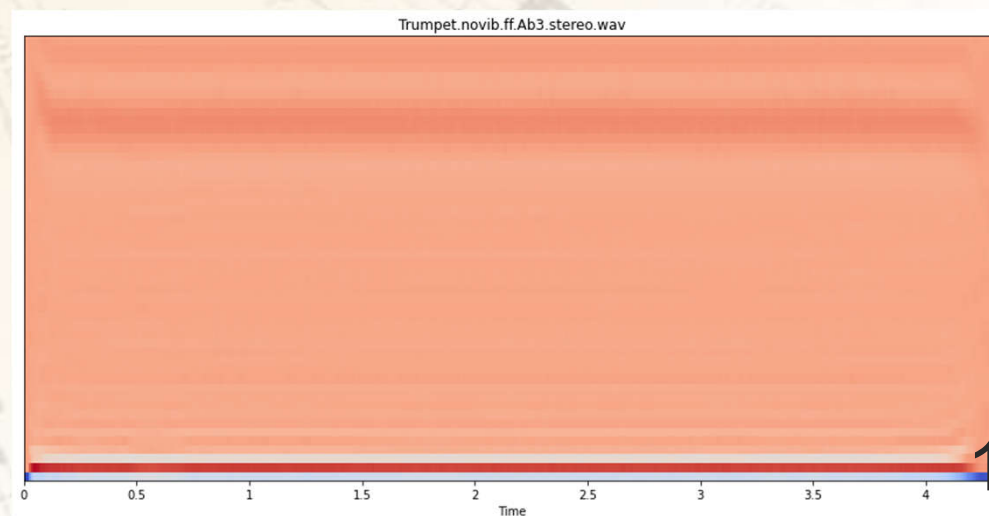
Скрипка



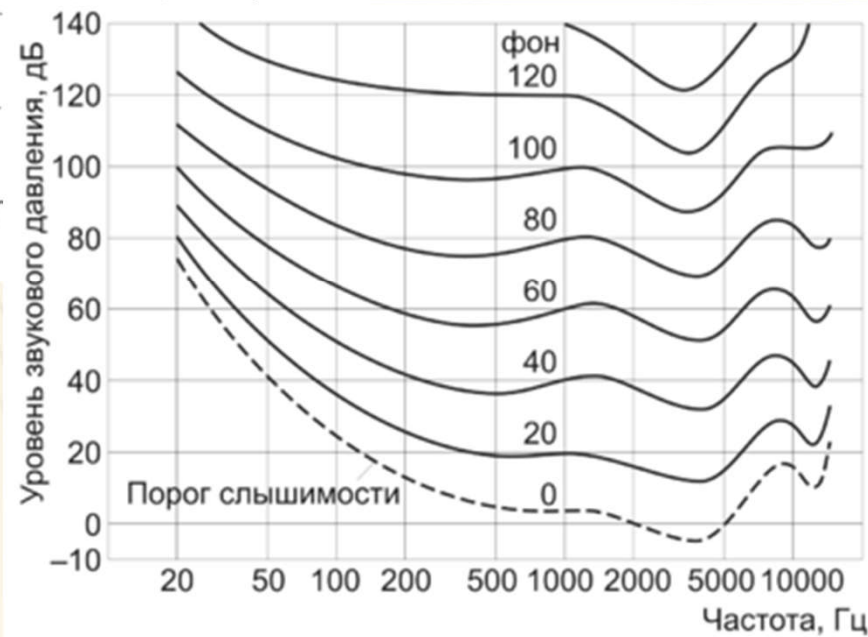
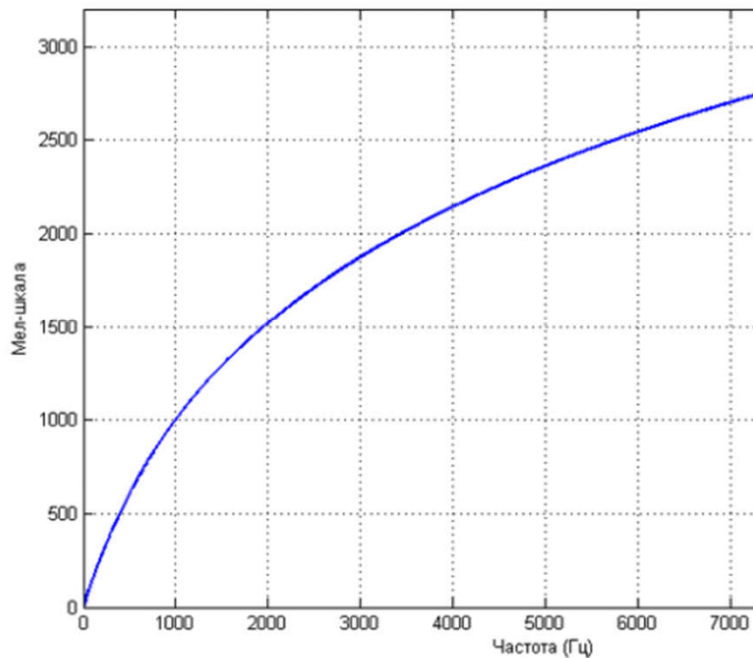
Гитара



Труба

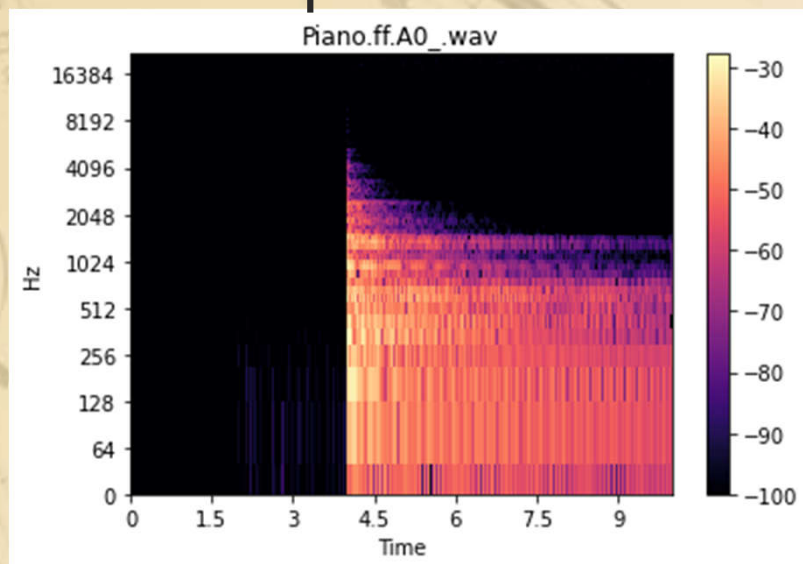


Мел-кепстральные коэффициенты. Мел-шкала

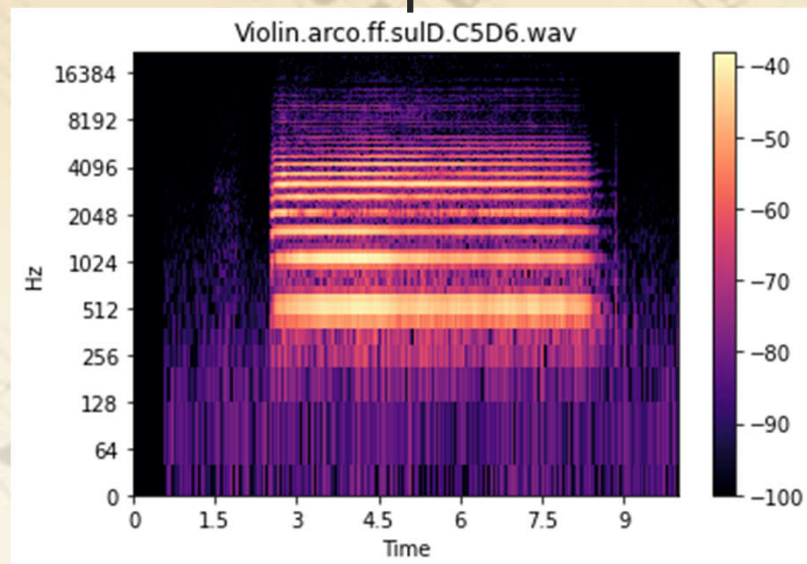


Мел-кепстральные коэффициенты. Преобразование Габора

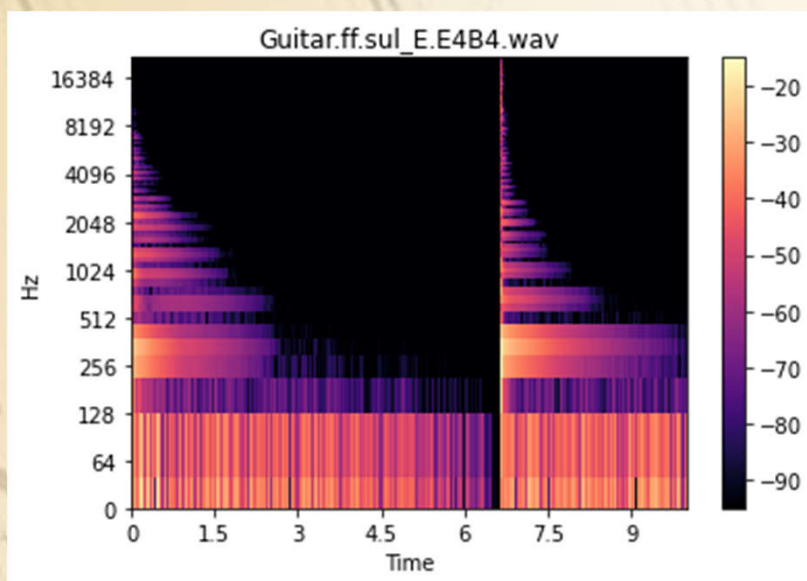
Фортепиано



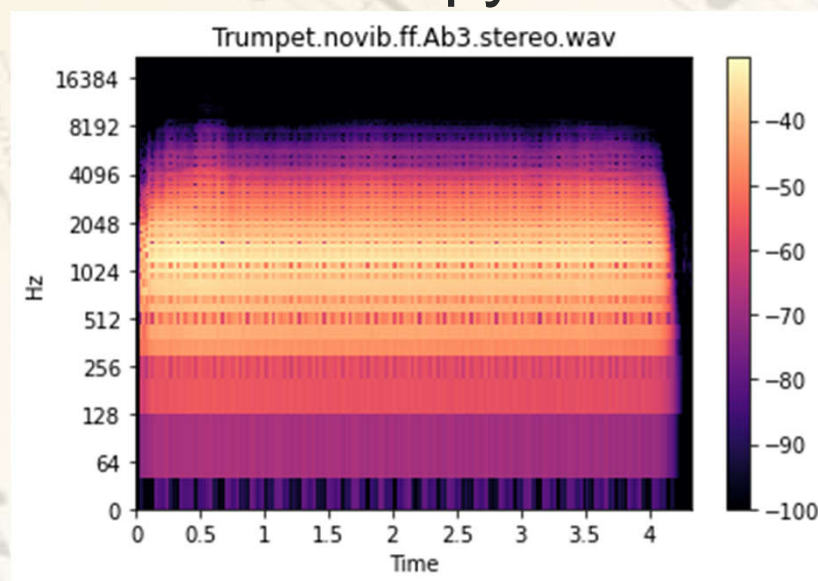
Скрипка



Гитара

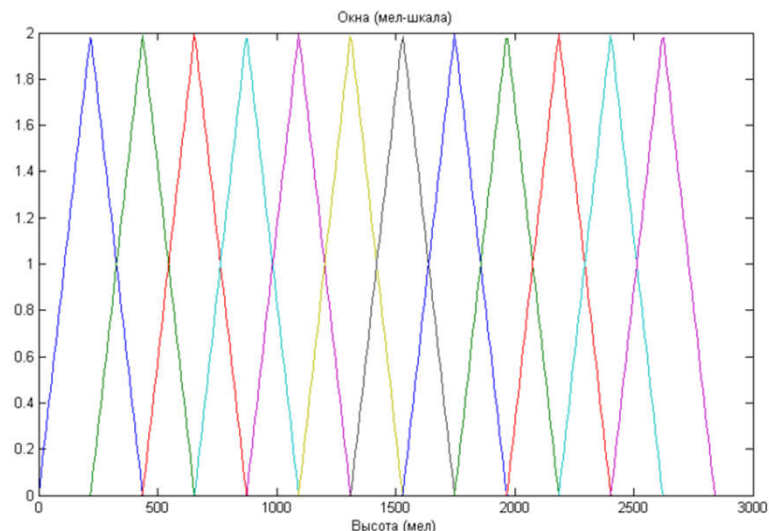


Труба

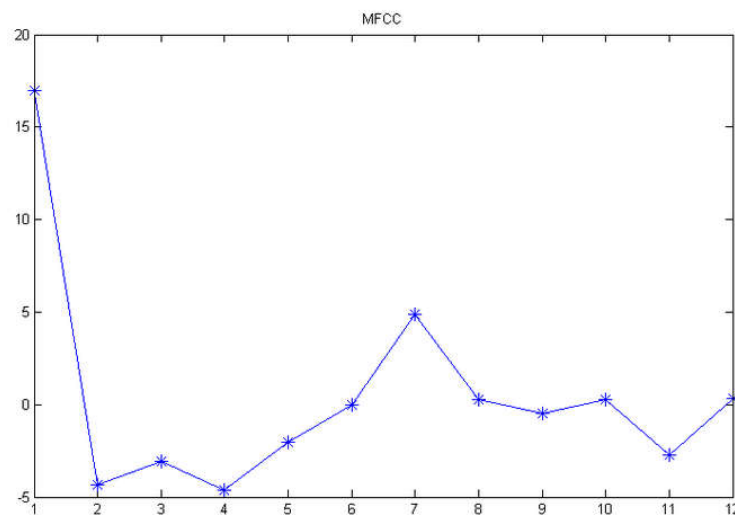


Мел-кепстральные коэффициенты.

Окна на мел-шкале

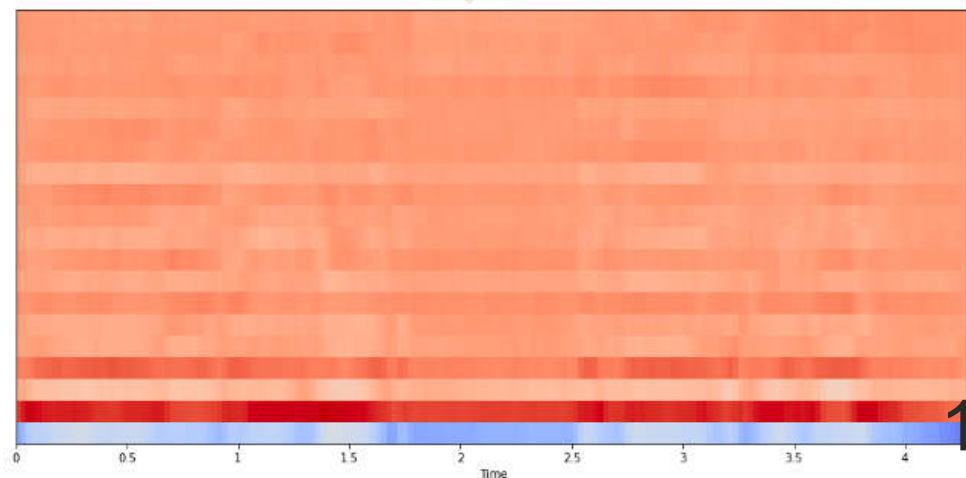


Результат преобразований:

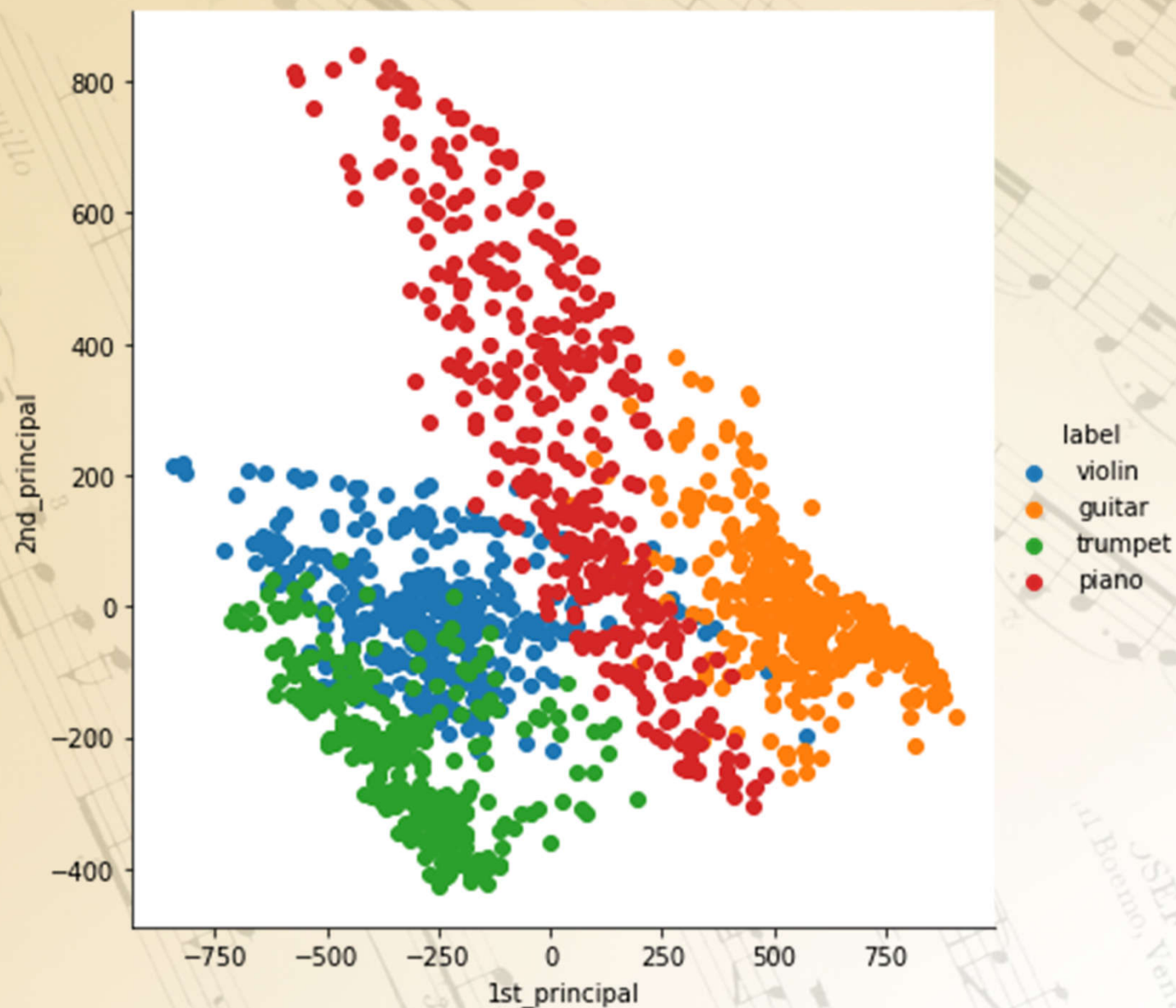


Результат для последовательности таймфреймов:

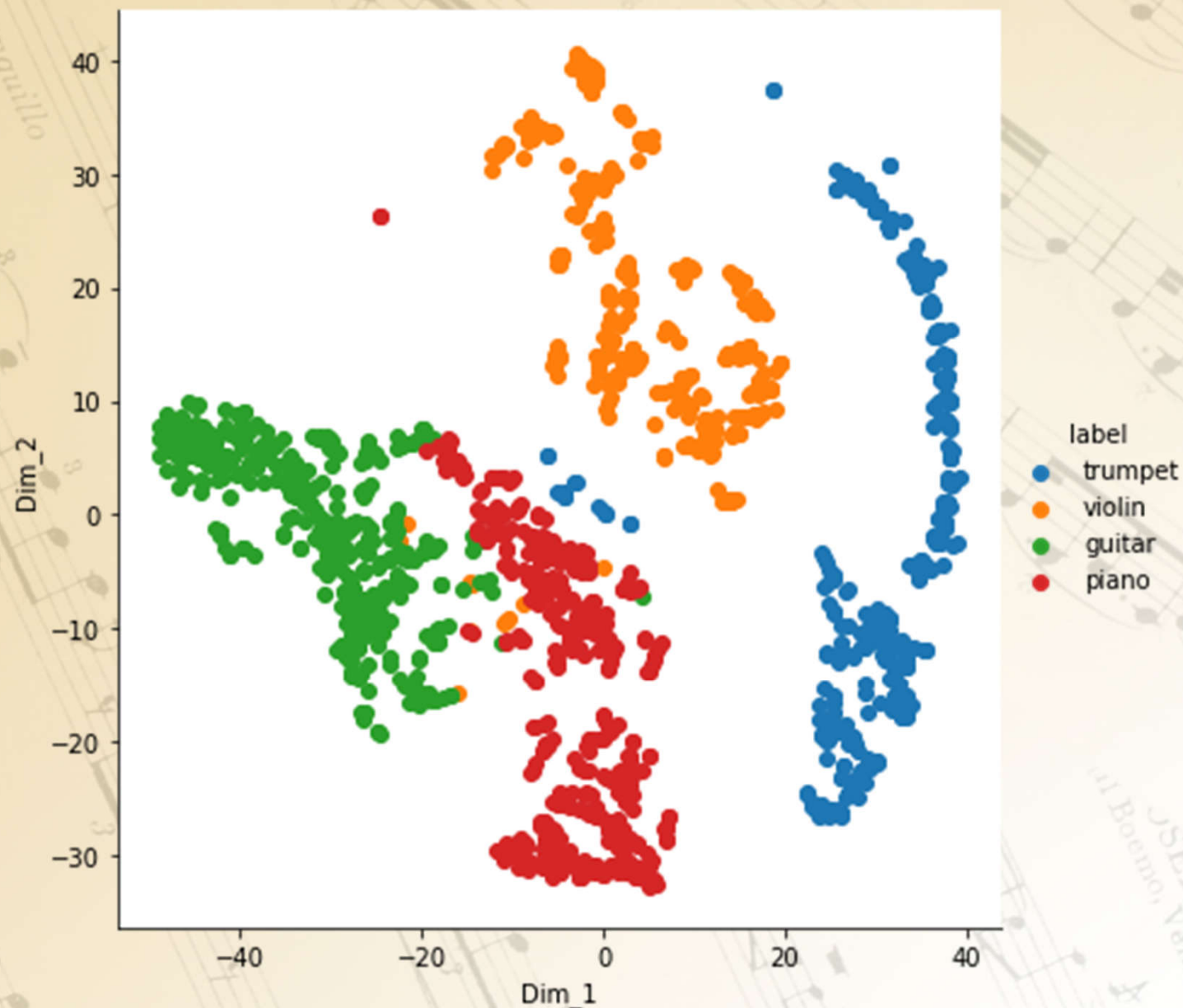
`librosa.feature.mfcc`



Визуализация признаков методом главных компонент



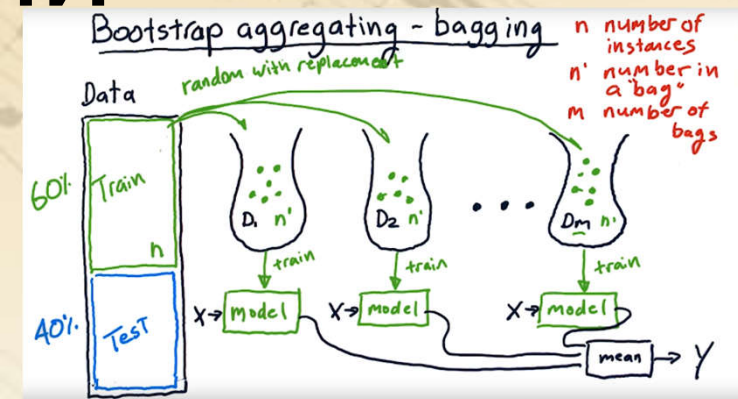
Визуализация признаков методом T-SNE



Модели классификации

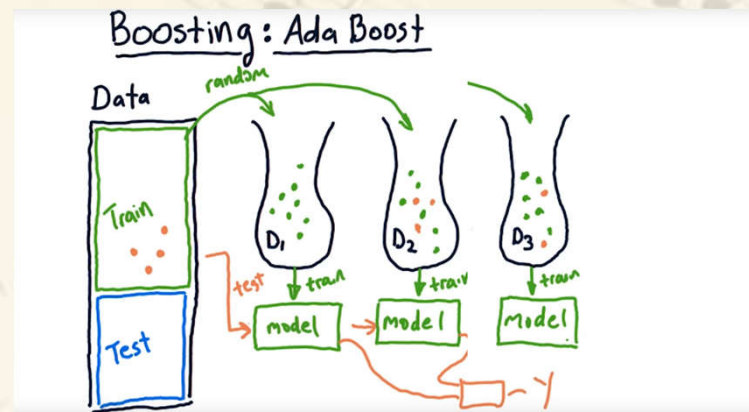
1. Случайный лес

- Использует модель бэггинга
- $n_estimators = 100$
- $max_depth = 10$
- Критерий разбиения – gini или entropy



2. Градиентный бустинг

- $n_estimators = 100$
- $max_depth = 10$
- $learning_rate = 0.1$



Результаты классификации методом RandomForest

Критерий разбиения — индекс Джини

```
Анализ датасета _22050_.csv gini
Обучающий набор score cross_val = 0.99
Тестовый набор score = 0.99
```

	piano	guitar	violin	trumpet
accuracy	0.99	1.00	0.96	1.0
precision	0.98	0.97	1.00	1.0
recall	0.99	1.00	0.96	1.0
f_measure	0.98	0.99	0.98	1.0

	piano	guitar	violin	trumpet
piano	79.0	0.0	2.0	0.0
guitar	1.0	73.0	1.0	0.0
violin	0.0	0.0	67.0	0.0
trumpet	0.0	0.0	0.0	64.0

$$\text{Gini}(Q) = 1 - \sum_{i=1}^n p_i^2,$$

p_i — вероятность i -го класса

Критерий разбиения — энтропия

```
Анализ датасета _22050_.csv entropy
Обучающий набор score cross_val = 0.99
Тестовый набор score = 0.97
```

	piano	guitar	violin	trumpet
accuracy	0.95	1.00	0.94	1.0
precision	0.98	0.90	1.00	1.0
recall	0.95	1.00	0.94	1.0
f_measure	0.96	0.95	0.97	1.0

	piano	guitar	violin	trumpet
piano	78.0	0.0	2.0	0.0
guitar	4.0	57.0	2.0	0.0
violin	0.0	0.0	64.0	0.0
trumpet	0.0	0.0	0.0	80.0

$$H = - \sum_{i=1}^n \frac{N_i}{N} \log \left(\frac{N_i}{N} \right)$$



Результаты классификации методом GradientBoosting

```
Анализ датасета _22050_.csv
Обучающий набор score cross_val = 0.89
Тестовый набор score = 0.89
```

	piano	guitar	violin	trumpet
accuracy	0.94	1.0	0.78	0.87
precision	0.81	1.0	0.86	0.89
recall	0.94	1.0	0.78	0.87
f_measure	0.87	1.0	0.82	0.88

	piano	guitar	violin	trumpet
piano	29.0	0.0	5.0	2.0
guitar	0.0	35.0	0.0	0.0
violin	2.0	0.0	31.0	3.0
trumpet	0.0	0.0	4.0	33.0

Также проводилось обучение модели с
обрезкой дерева по глубине.

Подробности на https://github.com/aleks1212v/data_mining





ЧАСТЬ 2

Выбор второго датасета - синтезированные инструменты

- Я выбрал для исследования синтезатор Casio с новейшим мощным музыкальным процессором AiX.
- Его сигнальный процессор позволяет синтезировать 600 музыкальных инструментов.
- Я выбрал 4 категории инструментов для исследования.



Исследуемые инструменты

PIANO

001	STAGE PIANO
002	GRAND PIANO
003	BRIGHT PIANO
004	MELLOW PIANO
005	POP PIANO
006	ROCK PIANO
007	DANCE PIANO
008	LA PIANO
009	TACK PIANO
010	MONO PIANO

STRINGS

192	PIZZICATO STRINGS
193	CHAMBER
194	OCTAVE STRINGS
195	ORCHESTRA HIT
196	SOLO VIOLIN
197	VIOLIN
198	SLOW VIOLIN
199	VIOLA
200	CELLO
201	CONTRABASS

GUITAR

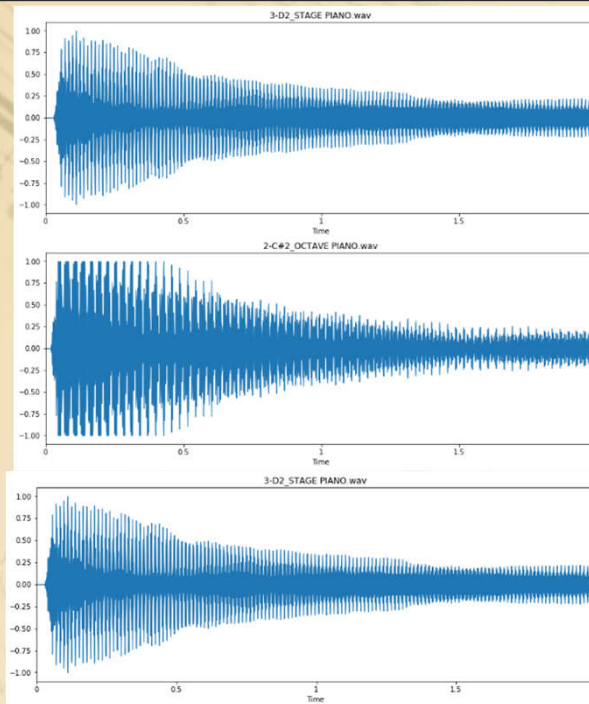
088	NYLON GUITAR VELO.SLIDE
089	STEEL GUITAR VELO.SLIDE
090	NYLON STR.GUITAR 1
091	NYLON STR.GUITAR 2
092	NYLON STR.GUITAR 3
093	STEEL STR.GUITAR 1
094	STEEL STR.GUITAR 2
095	STEEL STR.GUITAR 3
096	STEEL STR.GUITAR 4
097	12 STR.GUITAR

BRASS

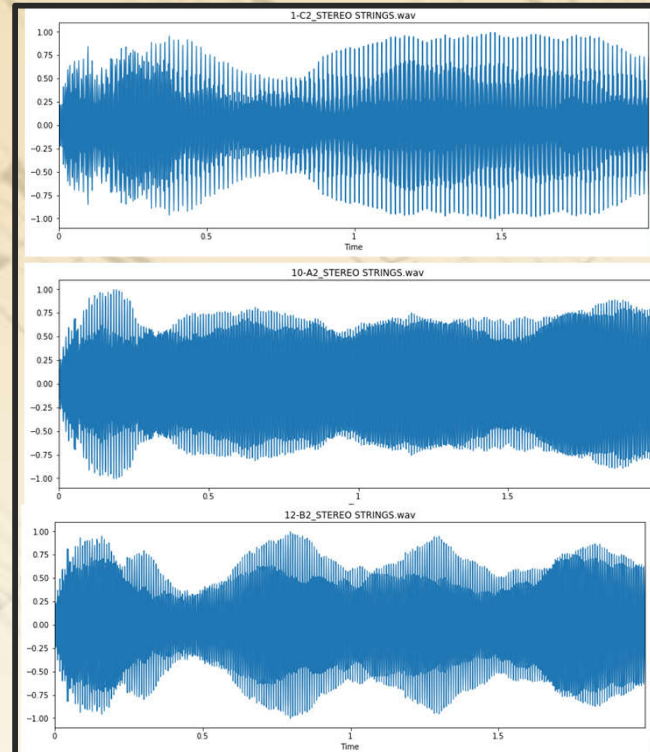
205	STEREO BRASS
206	SOLO TRUMPET
207	SYNTH-BRASS 1
208	BRASS
209	BRASS SECTION 1
210	BRASS SECTION 2
211	BIG BAND BRASS
212	HARD BRASS
213	BRASS SFZ
214	BRASS VELO.FALL



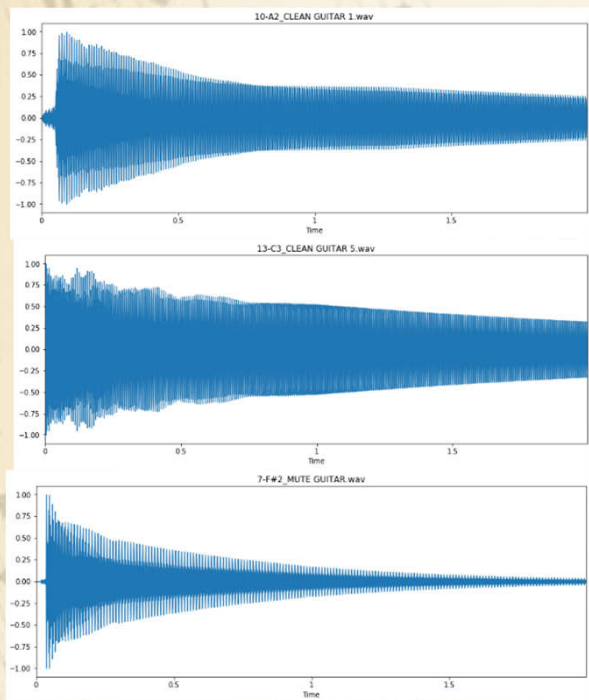
PIANO



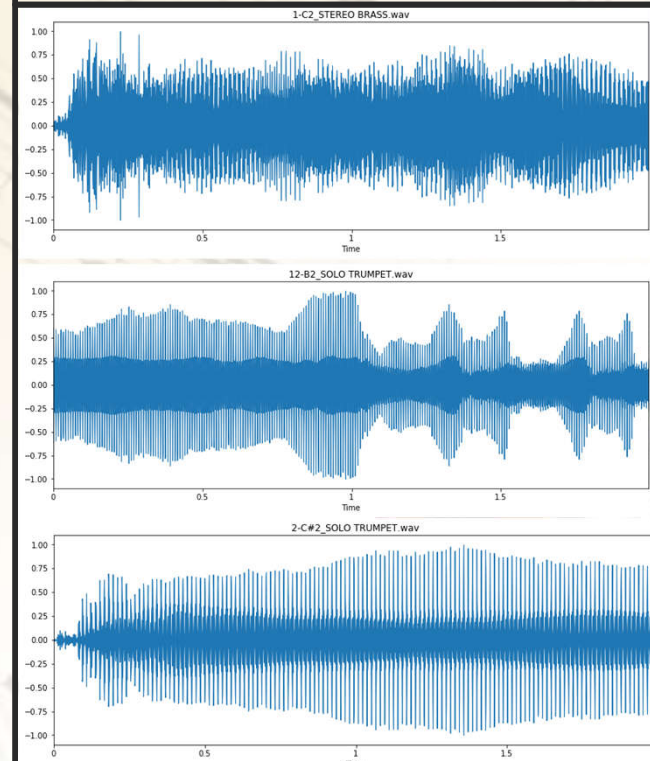
STRINGS



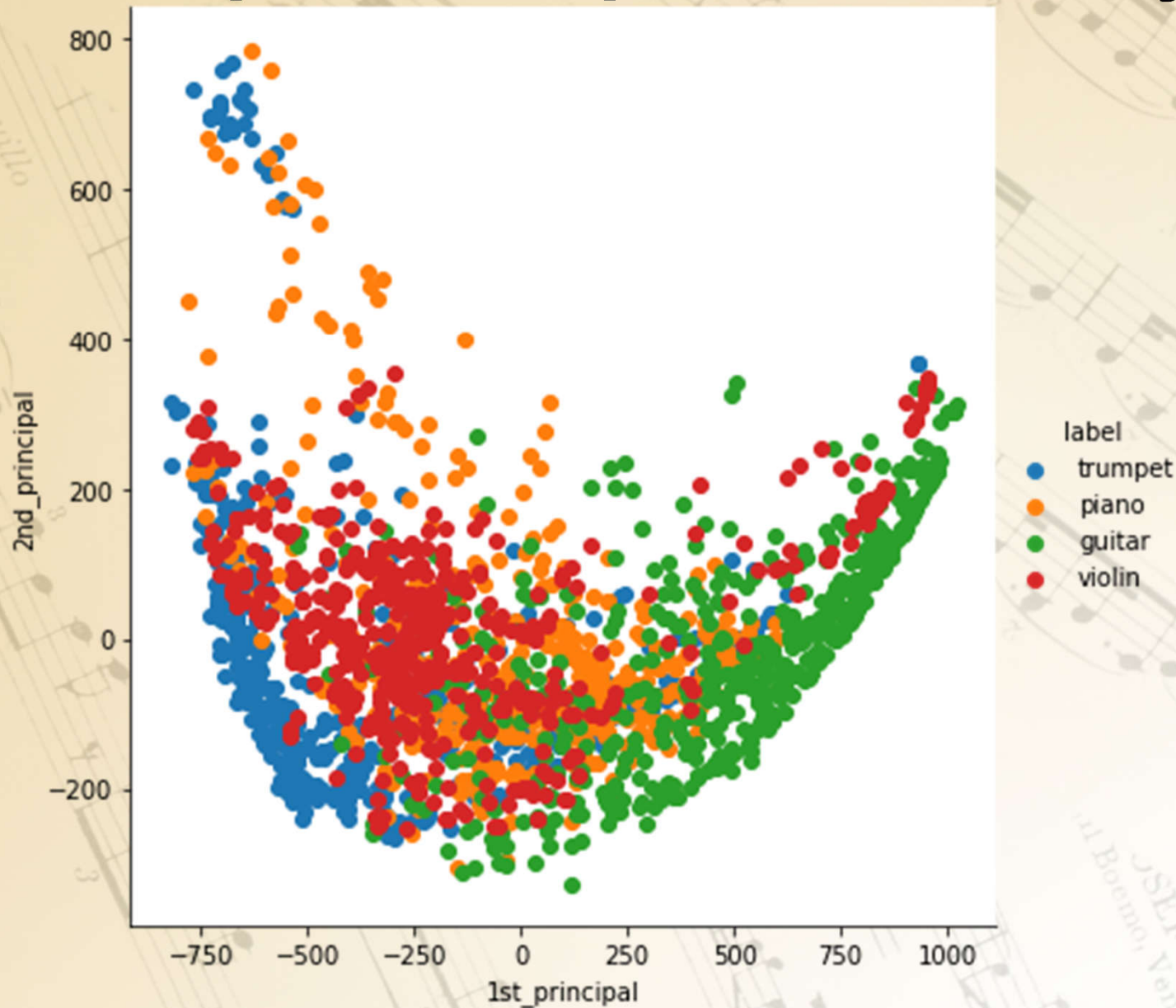
GUITAR



BRASS



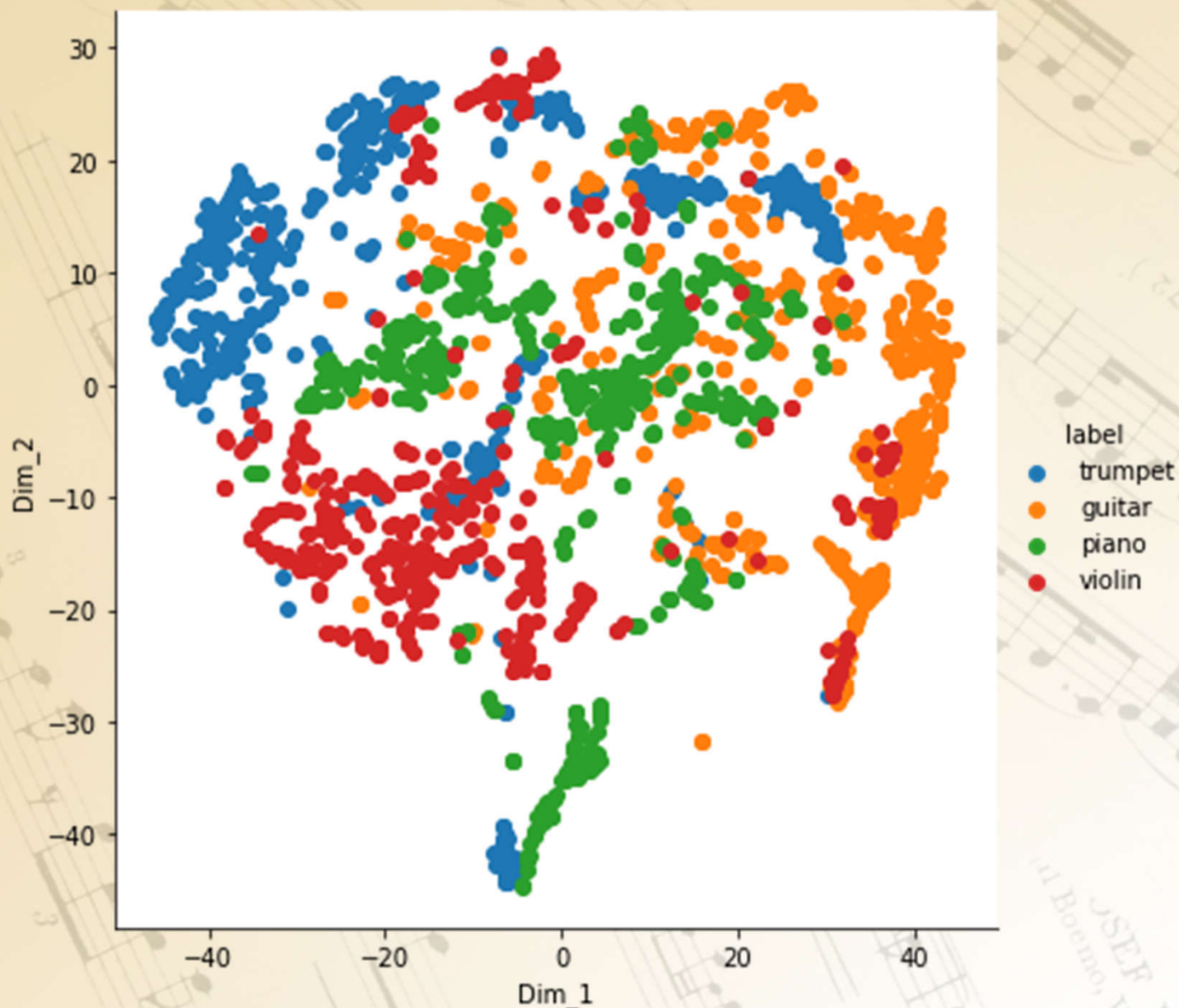
Визуализация РСА (principal component analysis)



github.com/aleks1212v/data_mining



Визуализация t-SNE



perplexity = 30, learning_rate = 10, n_iter = 1000



Результаты классификации

RandomForest

GradientBoosting

Критерий - индекс Джини

```
Анализ датасета _22050_.csv gini
Обучающий набор score cross_val = 0.92
Тестовый набор score = 0.93
```

	piano	guitar	violin	trumpet
accuracy	0.97	0.93	0.91	0.91
precision	0.90	0.94	0.92	0.97
recall	0.97	0.93	0.91	0.91
f_measure	0.93	0.94	0.91	0.94

	piano	guitar	violin	trumpet
piano	420.0	18.0	17.0	12.0
guitar	6.0	376.0	17.0	1.0
violin	4.0	7.0	404.0	26.0
trumpet	4.0	2.0	7.0	382.0

```
Анализ датасета _22050_.csv
Обучающий набор score cross_val = 0.81
Тестовый набор score = 0.82
```

	piano	guitar	violin	trumpet
accuracy	0.94	0.80	0.81	0.74
precision	0.94	0.74	0.79	0.82
recall	0.94	0.80	0.81	0.74
f_measure	0.94	0.77	0.80	0.78

	piano	guitar	violin	trumpet
piano	194.0	11.0	0.0	2.0
guitar	6.0	165.0	23.0	29.0
violin	2.0	13.0	167.0	30.0
trumpet	5.0	18.0	15.0	172.0

Критерий – энтропия

```
Анализ датасета _22050_.csv entropy
Обучающий набор score cross_val = 0.94
Тестовый набор score = 0.95
```

	piano	guitar	violin	trumpet
accuracy	0.98	0.94	0.94	0.93
precision	0.92	0.97	0.93	0.98
recall	0.98	0.94	0.94	0.93
f_measure	0.95	0.95	0.94	0.95

	piano	guitar	violin	trumpet
piano	419.0	21.0	12.0	5.0
guitar	5.0	398.0	9.0	0.0
violin	2.0	5.0	412.0	25.0
trumpet	3.0	1.0	3.0	383.0





ЧАСТЬ 3

Тестирование моделей на первом датасете (из Университета Айовы)

RandomForest

Критерий - индекс Джини

```
Проверочный набор score = 0.26
      piano guitar violin trumpet
accuracy  0.21  0.08  0.03  0.73
precision 0.21  0.86  0.05  0.32
recall    0.21  0.08  0.03  0.73
f_measure 0.21  0.15  0.04  0.44

      piano guitar violin trumpet
piano    74.0  170.0  30.0  79.0
guitar    1.0   30.0   1.0   3.0
violin   41.0  156.0  11.0  16.0
trumpet 242.0   2.0  316.0 260.0
```

Критерий – энтропия

```
Проверочный набор score = 0.26
      piano guitar violin trumpet
accuracy  0.13  0.06  0.03  0.82
precision 0.17  0.62  0.04  0.33
recall    0.13  0.06  0.03  0.82
f_measure 0.15  0.12  0.03  0.48

      piano guitar violin trumpet
piano    46.0  157.0  23.0  47.0
guitar    3.0   23.0   0.0  11.0
violin   54.0  176.0  10.0   7.0
trumpet 255.0   2.0  325.0 293.0
```

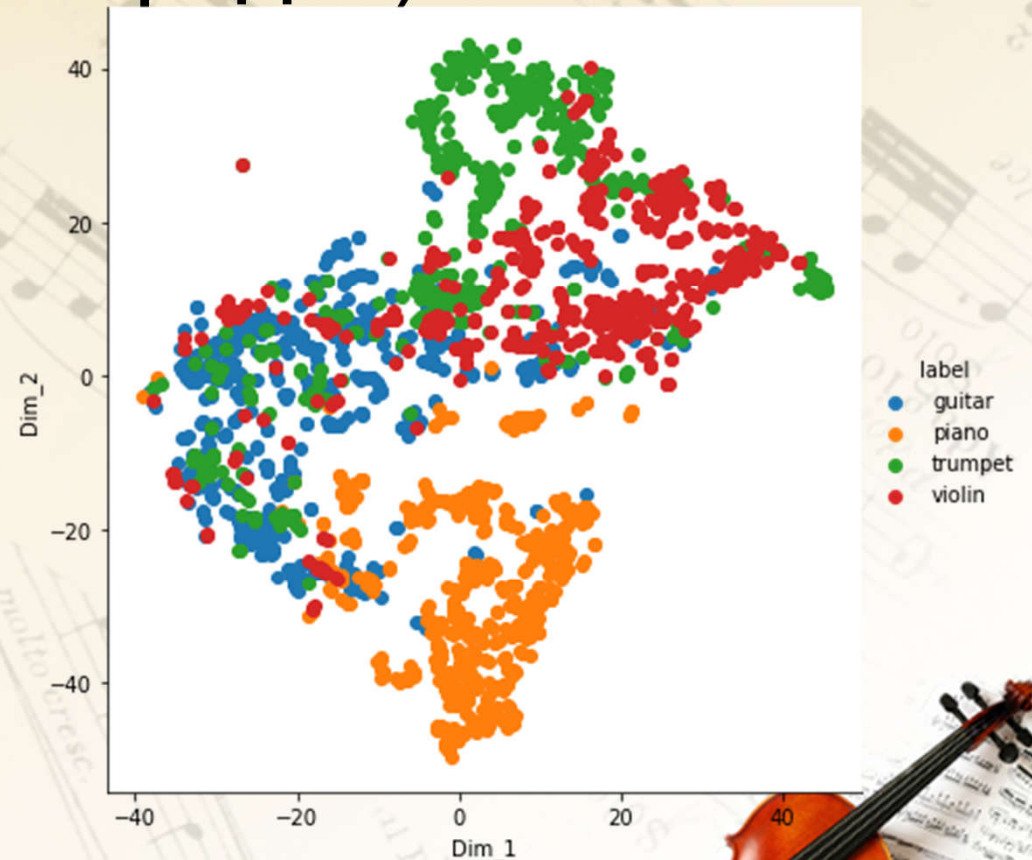
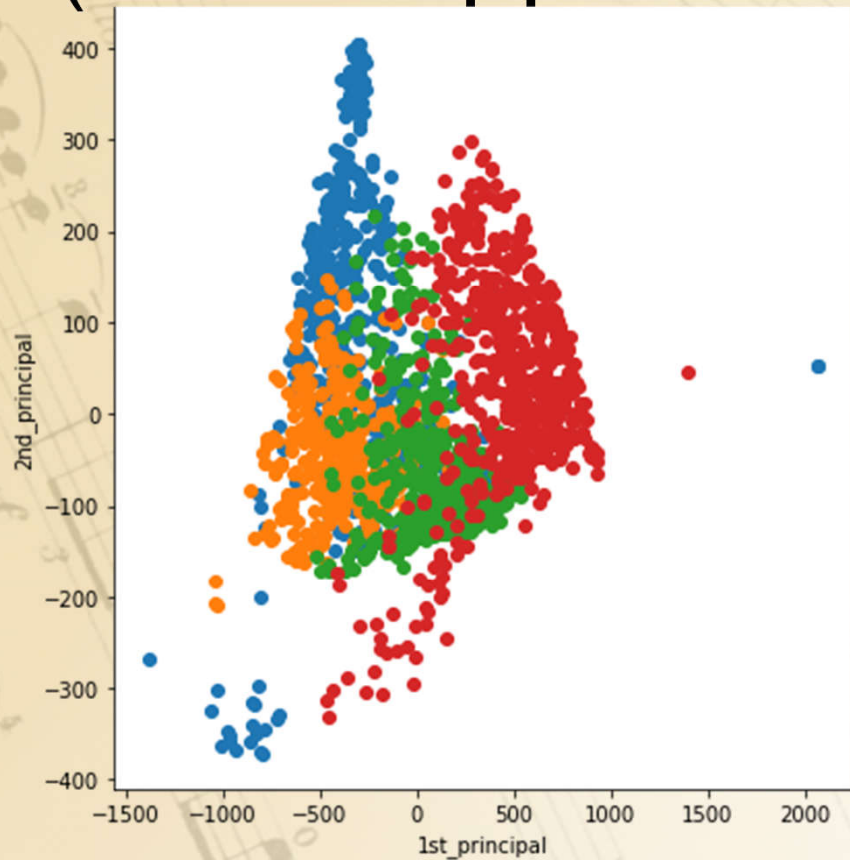
GradientBoosting

```
Проверочный набор score = 0.2
      piano guitar violin trumpet
accuracy  0.03  0.01  0.05  0.71
precision 0.62  0.01  0.08  0.29
recall    0.03  0.01  0.05  0.71
f_measure 0.05  0.01  0.06  0.41

      piano guitar violin trumpet
piano    10.0   0.0   5.0   1.0
guitar   23.0   2.0  213.0  57.0
violin   33.0  142.0  18.0  45.0
trumpet 292.0  214.0  122.0 255.0
```



Визуализация признаков после выделения особенностей путем вейвлет-преобразования исходного сигнала (вейвлет Добеши 2-го порядка)



Результаты классификации после выделения особенностей сигнала

Модель RandomForest

Критерий - индекс Джини

Анализ датасета _22050_db2_cD.csv gini
Обучающий набор score cross_val = 0.87
Тестовый набор score = 0.88

	piano	guitar	violin	trumpet
accuracy	0.96	0.91	0.87	0.77
precision	0.97	0.77	0.90	0.90
recall	0.96	0.91	0.87	0.77
f_measure	0.97	0.83	0.88	0.83

	piano	guitar	violin	trumpet
piano	210.0	5.0	1.0	1.0
guitar	8.0	196.0	18.0	32.0
violin	0.0	7.0	186.0	14.0
trumpet	0.0	8.0	9.0	157.0

Проверочный набор score = 0.49

	piano	guitar	violin	trumpet
accuracy	0.0	0.91	0.06	0.99
precision	0.0	0.95	0.09	0.41
recall	0.0	0.91	0.06	0.99
f_measure	0.0	0.93	0.07	0.59

	piano	guitar	violin	trumpet
piano	0.0	1.0	10.0	0.0
guitar	0.0	162.0	8.0	0.0
violin	94.0	1.0	10.0	1.0
trumpet	85.0	15.0	151.0	178.0

Критерий – энтропия

Анализ датасета _22050_db2_cD.csv entropy
Обучающий набор score cross_val = 0.88
Тестовый набор score = 0.88

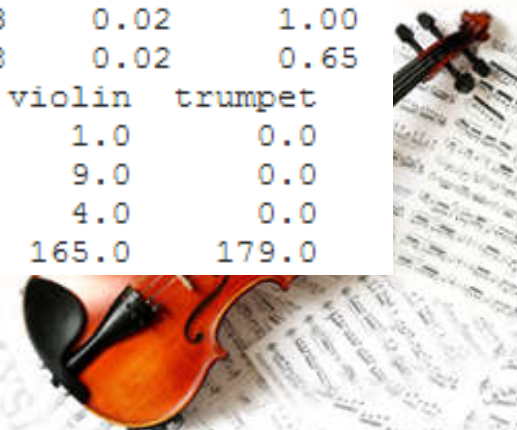
	piano	guitar	violin	trumpet
accuracy	0.99	0.94	0.86	0.75
precision	0.98	0.79	0.86	0.92
recall	0.99	0.94	0.86	0.75
f_measure	0.98	0.85	0.86	0.82

	piano	guitar	violin	trumpet
piano	198.0	4.0	1.0	0.0
guitar	2.0	187.0	22.0	27.0
violin	1.0	6.0	204.0	27.0
trumpet	0.0	3.0	11.0	159.0

Проверочный набор score = 0.46

	piano	guitar	violin	trumpet
accuracy	0.0	0.83	0.02	1.00
precision	0.0	0.93	0.02	0.48
recall	0.0	0.83	0.02	1.00
f_measure	0.0	0.88	0.02	0.65

	piano	guitar	violin	trumpet
piano	0.0	4.0	1.0	0.0
guitar	2.0	148.0	9.0	0.0
violin	168.0	4.0	4.0	0.0
trumpet	9.0	23.0	165.0	179.0



БИБЛИОГРАФИЯ

1. Станкевич Ф.В., Спицын В.Г. НЕЙРОСЕТЕВОЕ РАСПОЗНАВАНИЕ МУЗЫКАЛЬНЫХ ИНСТРУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ МЕЛ-ЧАСТОТНЫХ КЕПСТРАЛЬНЫХ КОЭФФИЦИЕНТОВ. ФГАОУ ВО «Национальный исследовательский Томский политехнический университет», Томск
2. Орельен Жерон. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow. Концепции, инструменты и техники для создания интеллектуальных систем. / Орельен Жерон. –Москва: Вильямс, 2018. –688 с.
3. Захарова, Т.В. Вейвлет-анализ и его приложения. Учебное пособие / Т.В. Захарова, О.В. Шестаков. Второе издание, переработанное и дополненное. Москва: ИНФРА-М, 2019. –159 с.[Высшее образование]

