

Классификация музыки по жанрам

Проект по машинному обучению

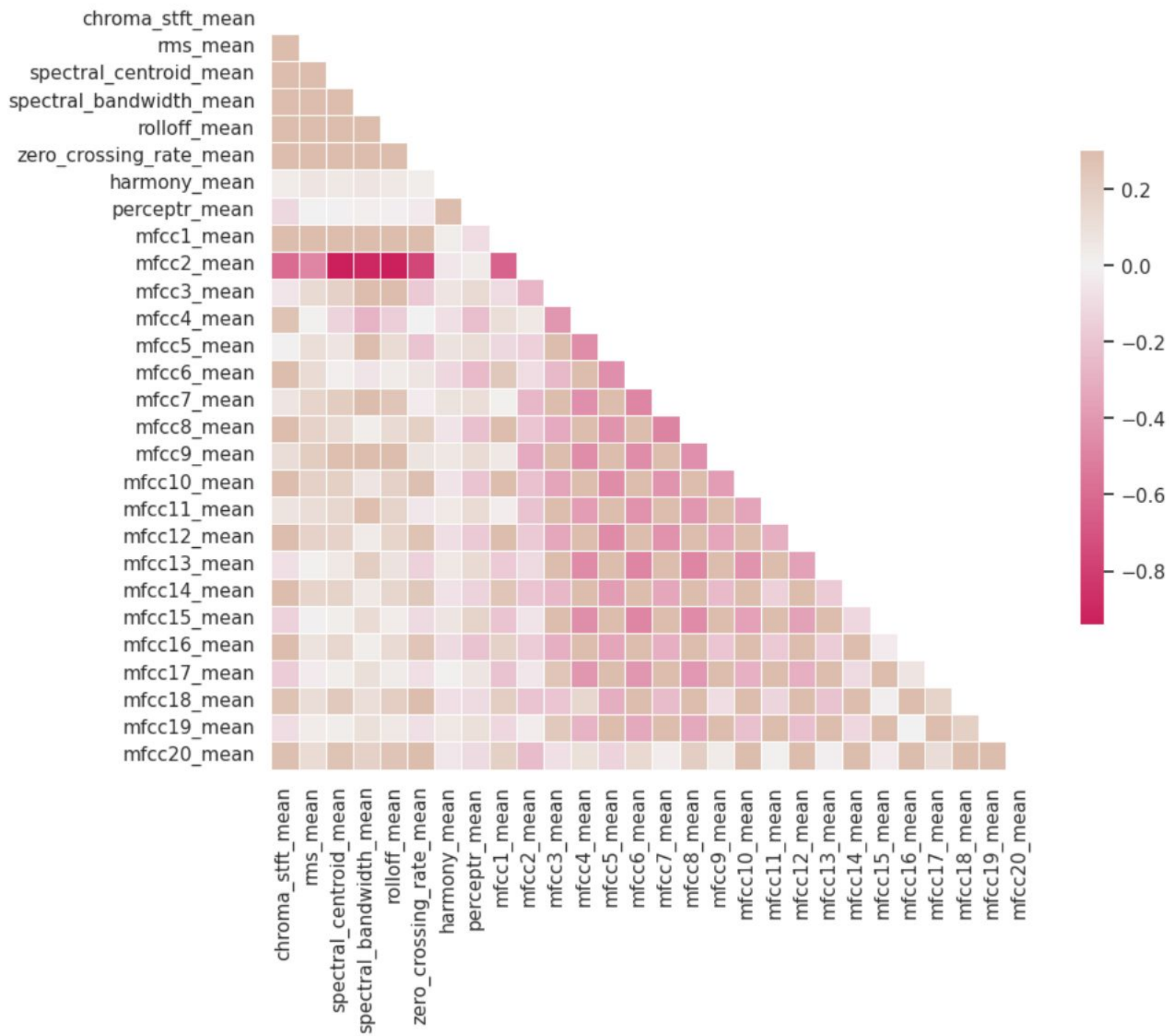


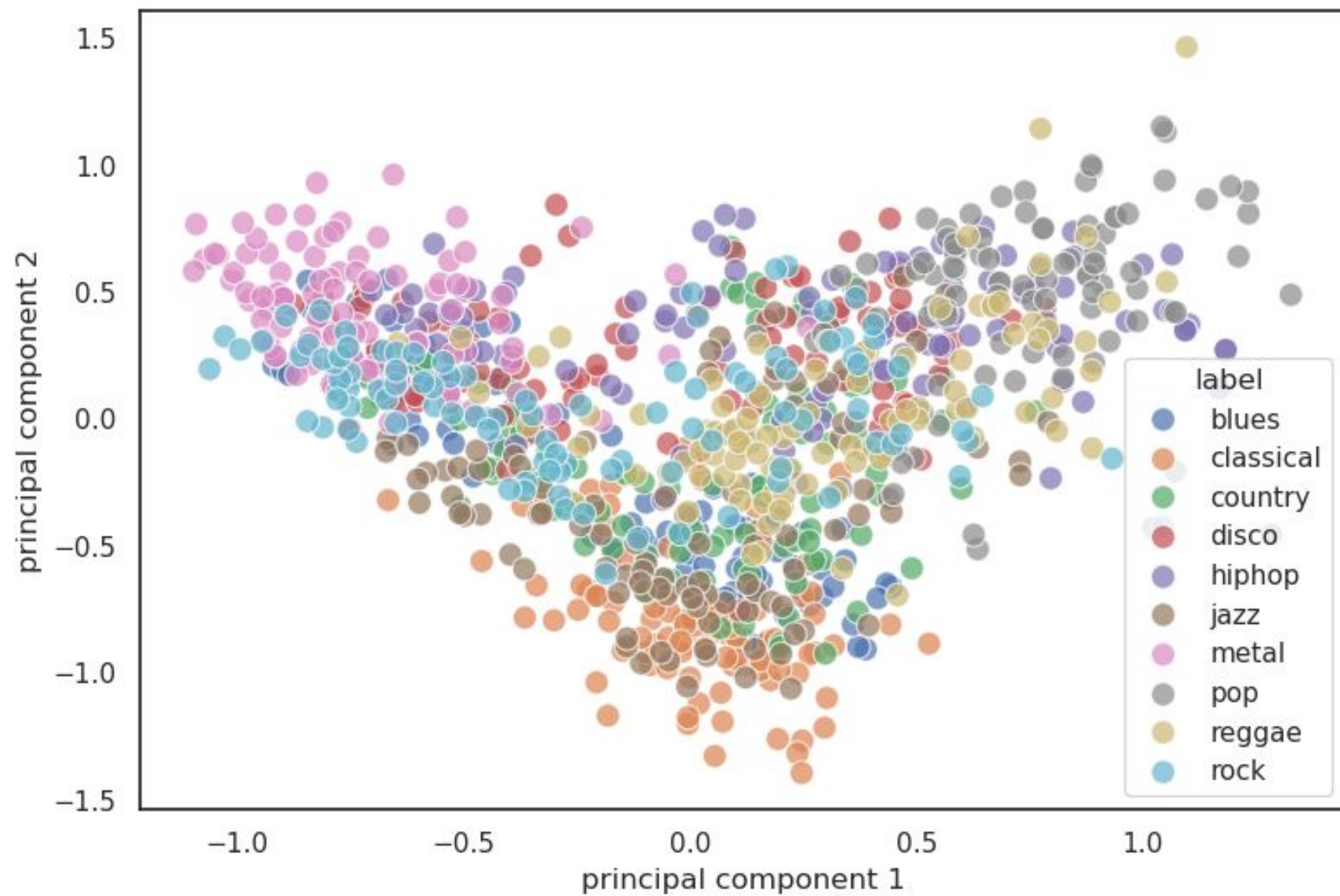
Цель работы

Научиться
классифицировать
аудиозаписи
музыкального фрагмента
по жанрам.

Исходные данные

- 1000 аудиофайлов по 30 секунд
- Те же файлы, нарезанные на маленькие кусочки - 10 000 файлов по 3 секунды
- 10 сбалансированных классов
- Фичи - средние и дисперсии различных спектральных данных, гармонических и ударных сигналов и т.д.
- Картинки со спектрограммами





Используемые модели

Модель	f1_score
LogisticRegression	0.68
RandomForestClassifier	0.85
KNeighborsClassifier	0.88
CatBoostClassifier	0.90
SVC	0.74
XGBClassifier	0.88

Итоговая модель

- train, val, test выборки (соотношение размера 6:2:2)
- MinMaxScaler
- С помощью optuna подобрали параметры CatBoost
- Итоговый f1 score = **0.92**

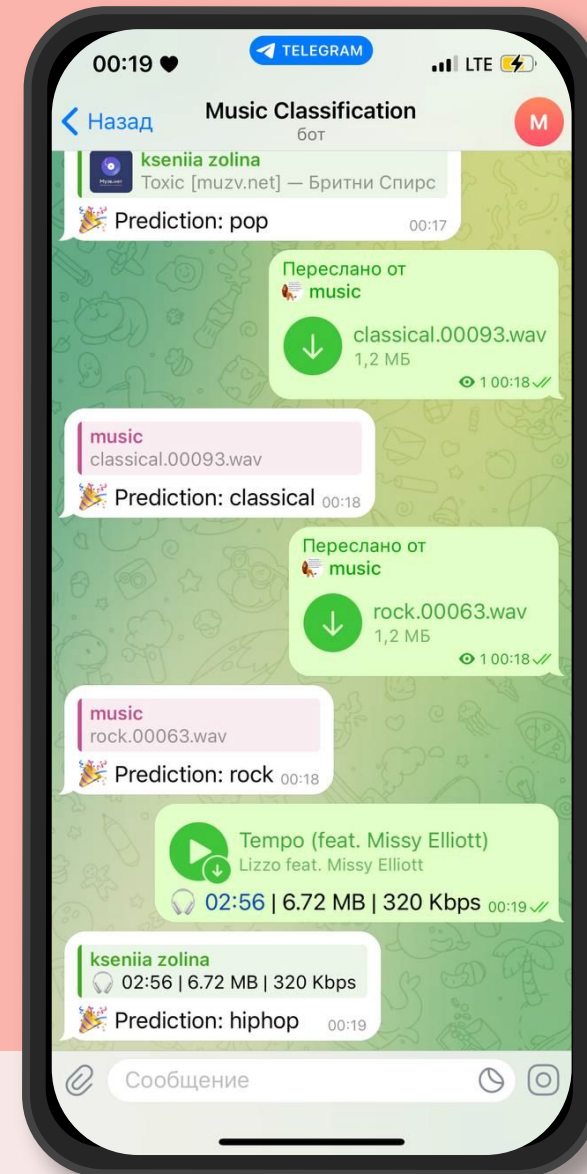
А что там с картинками?

- 3 сверточных слоя
- 3 линейных слоя
- BatchNorm
- 6 эпох
- CrossEntropyLoss

...f1 score = 0.19

Мы полагаем, что данных слишком мало - в обучающей выборке всего 600 спектрограмм.

Демо бота



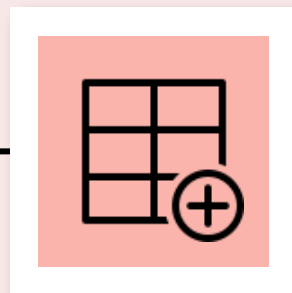
Пайплайн бота

**Принимает .wav
или .mp3 файл**



Извлекает фичи

С помощью библиотеки
librosa



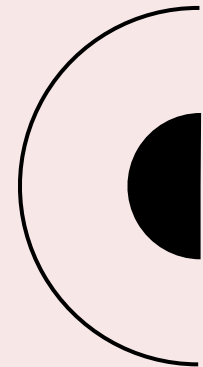
**Выбирает самый
частый класс**



**Разбивает на
короткие куски**

Длина куска - 3 секунды,
как и была при обучении

**Предсказывает
для каждого
куска**



Итоговые результаты

- f1_score (CatBoostClassifier): 0.903105 → **0.922965**
(тюнинг с помощью optuna)
- Научились извлекать фичи из любого аудио (?)
- Модель обернута в телеграм-бота



**Спасибо за
ВНИМАНИЕ!**