

РЕТ-проект "Определение жанра по изображению обложки музыкального диска"

Описание проекта

Рет-проект, чтобы потренировать навыки, приобрести новые и нескучно провести время.

Очевидно, что оформление музыкального альбома как-то связано с его содержанием. Но насколько связано? Как подтвердить это, опираясь на данные? И чем это может быть полезно?

Основная задача

Необходимо разработать модель, которая определит жанр музыкального альбома по изображению его обложки.

Дополнительные задачи

1. **Кластеризация:** исследовать, какие жанры лучше выделяются на фоне других, а какие – трудноразделимы.
2. **API [MusicBrainz](#):** расширение датасета новыми изображениями. Можно расширить список жанров и использовать обратную сторону обложки.
3. **OCR + NLP:** извлечь текст из изображения и обогатить решение применением NLP.
4. Создать **рекомендательную систему** в миниатюре, опираясь на изображения.
5. **Streamlit** оформить классификатор жанров или рекомендательную систему в виде веб-приложения.

Данные

Изображения в формате PNG, упакованные в zip-архивы. Имя архива соответствует музыкальному жанру. Каждый zip-архив содержит папку с изображениями музыкальных обложек соответствующего жанра.

Метрика

Метрики для классификации: Accuracy, Precision, Recall, F1. Можно предложить свои метрики как часть исследования.

Рекомендуемый план реализации

1. Ознакомление с данными.
2. Предварительная обработка данных.
3. Определение целевой метрики.
4. Подготовка данных для baseline-модели.
5. Создание baseline-модели.
6. Расчет метрик.
7. Подготовка данных для нейронной сети.
8. Обучение нейронной сети.
9. Расчет и сравнение метрик.
10. Анализ финальной модели.
11. Исследование/решение дополнительных задач.
12. Подготовка отчета по исследованию.

Оформление

1. Тетрадь с решением задачи.
2. Если решение состоит из нескольких частей — репозиторий на GitHub.
3. Если сделано веб-приложение — репозиторий на GitHub и ссылка на приложение.

Срок реализации

Срок реализации проекта - 2 недели.

1-я неделя: изучение данных, создание baseline.

2-я неделя: обучение нейронной сети.

Рекомендуемый стек

python, sklearn, faiss, pytorch, torchvision, fastai.

Полезные материалы

1. [MusicBrainz artwork downloader](#)
2. [Python-musicbrainzngs](#)
3. [FAISS: Быстрый поиск лиц и клонов на многомиллионных данных](#)
4. [Introduction to FAISS](#)
5. [Fine-Tuning ResNet50](#)
6. [Глубокое обучение быстро при помощи Fast AI](#)
7. [Fast AI docs](#)
8. [Fast AI github](#)