Яндекс Медиасервисы

Распознавание музыки в кино и сериалах

КИНОПОИСК

О себе

Лёша Царёв — аналитик в Службе Развития Технологий Яндекс

О команде

Служба развития технологий
Рекомендации
Медиапоиск
Голосовые сценарии
Исследование технологий

О команде

Служба развития технологий
Крауд
Рекомендации
Медиапоиск
Голосовые сценарии
Исследование технологий
АБ-эксперименты

Что такое краудсорсинг?

Что такое краудсорсинг?

О команде

Метрики
Служба развития технологий
Крауд
Рекомендации
Медиапоиск
Голосовые сценарии
Исследование технологий
АБ-эксперименты

Чем занимается наша команда?

- Музыкальные рекомендации
- Рекомендации фильмов и сериалов
- Распознавание музыки и актёров в кино и сериалах
- Плейлист с Алисой
- Генеративная музыка

Чем занимается наша команда?

- Музыкальные рекомендации
- Рекомендации фильмов и сериалов
- Распознавание музыки и актёров в кино и сериалах
- Плейлист с Алисой
- Генеративная музыка

Чем занимается наша команда?

- Музыкальные рекомендации
- Рекомендации фильмов и сериалов
- Распознавание музыки и актёров в кино и сериалах
- Плейлист с Алисой
- Генеративная музыка

Чем занимается наша команда?

- Музыкальные рекомендации
- Рекомендации фильмов и сериалов
- Распознавание музыки и актёров в кино и сериалах
- Плейлист с Алисой
- Генеративная музыка

Чем занимается наша команда?

- Музыкальные рекомендации
- Рекомендации фильмов и сериалов
- Распознавание музыки и актёров в кино и сериалах
- Плейлист с Алисой
- Генеративная музыка

Agenda

Что? Как? Зачем?

Кросс-сервисность

Развитие технологий

Фич-реквест

Постановка задачи

КИНОПОИСК

Как устроен процесс?

Информация о фильме и о музыке в фильме

Кинопоиск

Яндекс Музыка

Редакторская разметка

Как устроен процесс?

Находим музыку

Качество музыкального классификатора

98%

полнота

86%

точность

Проверяем классификатор

Как устроен процесс?

Находим музыку

Характеристики алгоритмов распознавания

	Алгоритм №1	Алгоритм №2
Алгоритм	ML-ный	Эвристический
Размер индекса	8 млн	1.4 млн
Акустика	«Уличная»	«Киношная»

Проверяем распознавание

Как устроен процесс?

Мёрджим результаты

Как устроен процесс?

Как тестировали качество?

Тестовый стенд

Метрики

Полнота разметки:

- Доля размеченного tvt, уников, каталога

Качество разметки:

- Доля распознанной музыки
- Доля распознанных саундтреков

Проблемные случаи:

- Доля фильмов, в которых распознали хотя бы 10% саундтреков

Метрики

Полнота разметки:

- Доля размеченного tvt, уников, каталога

Качество разметки:

- Доля распознанной музыки
- Доля распознанных саундтреков

Проблемные случаи:

- Доля фильмов, в которых распознали хотя бы 10% саундтреков

Метрики

Полнота разметки:

- Доля размеченного tvt, уников, каталога

Качество разметки:

- Доля распознанной музыки
- Доля распознанных саундтреков

Проблемные случаи:

- Доля фильмов, в которых распознали хотя бы 10% саундтреков

Какие проблемы были и как их решали?

Проблема:

Недостаточно информации о треках в фильмах

Решение:

Пополнять информацию
при помощи редакторов

Итого:

+10 п.п. к доле
распознанного времени

Было/Стало

Autoscroll
music
music
music
music
music
music
auto_music
music
00
05:00
10:00
15:00
20:00
25:00
30:00
35:00
40:00
45:00
50:00
55:00
1:00:00
1:05:00
1:10:00
1:15:00
1:20:00
1:25:00
Back

Autoscroll
music
music
music
music
music
music
auto_music
music
00
05:00
10:00
15:00
20:00
25:00
30:00
35:00
40:00
45:00
50:00
55:00
1:00:00
1:05:00
1:10:00
1:15:00
1:20:00
1:25:00
Back

Проблема:

24 fps vs 25 fps

Решение:

Замедляем аудио на 4%

Итого:

+13 п.п. к доле фильмов, в которых распознали хотя бы 10% саундтреков

Было/Стало

Autoscroll
auto_music
music
00
05:00
10:00
15:00
20:00
25:00
30:00
35:00
40:00
45:00
50:00
55:00
1:00:00
1:05:00
1:10:00
1:15:00
1:20:00
1:25:00
1:30:00
1:35:00
1:40:00
1:45:00
Back

Autoscroll
auto_music
music
00
05:00
10:00
15:00
20:00
25:00
30:00
35:00
40:00
45:00
50:00
55:00
1:00:00
1:05:00
1:10:00
1:15:00
1:20:00
1:25:00
1:30:00
1:35:00
1:40:00
1:45:00
Back

Проблема:

В некоторых треках содержатся куски из фильмов

Решение:

Сделать механизмы ручного вмешательства в разметку

Проблема:

Распознаём эксплисит треки в детском контенте

Решение:

В контенте до 12+ баним все эксплисит треки

Как ещё улучшали качество?

Улучшения

- Автоматизированная помощь в поиске саундтреков для редакторов
- Коррекция границ прогнозов распознанных треков
- Работа с очень маленькими распознанными фрагментами
- Обмениваться треками между алгоритмами распознавания

Улучшения

- Автоматизированная помощь в поиске саундтреков для редакторов
- Коррекция границ прогнозов распознанных треков
- Работа с очень маленькими распознанными фрагментами
- Обмениваться треками между алгоритмами распознавания

Улучшения

- Автоматизированная помощь в поиске саундтреков для редакторов
- Коррекция границ прогнозов распознанных треков
- Работа с очень маленькими распознанными фрагментами
- Обмениваться треками между алгоритмами распознавания

Улучшения

- Автоматизированная помощь в поиске саундтреков для редакторов
- Коррекция границ прогнозов распознанных треков
- Работа с очень маленькими распознанными фрагментами
- Обмениваться треками между алгоритмами распознавания

Сравнительный вклад в улучшение

Метрика: Распознанный TLT

Что получили по итогу?

Сейчас играет

O Children
Nick Cave & The Bad Seeds

Яндекс Музыка | Открыть | Сохранить

Актёры в кадре

Дэниэл Рэдклифф
Harry Potter

Эмма Уотсон
Hermione Granger

−01:01:42

Аудио и субтитры

Disney
Круэлла
ОСКАР 2022
Сейчас играет
Hush
Deep Purple
Яндекс Музыка
Открыть
Сохранить
Актёры в кадре
Марк Стронг
John the Valet
Эмма Томпсон
The Baroness
−01:24:21
Аудио и субтитры

Властелин Колец
Возвращение короля
Сейчас играет
The Battle of the Pelennor Fie
Howard Shore
Яндекс Музыка
Открыть
Сохранить
-01:32:56
Аудио и субтитры

Полученные результаты

Покрытие фичёй

- 93% пользователей увидели фичу
- На 75% того, что они смотрят

Полнота покрытия распознавания

- 84% на чистых фрагментах
- 46% на шумных фрагментов
- 17% на фоновой музыке

Полученные результаты

Покрытие фичёй

- 93% пользователей увидели фичу
- На 75% того, что они смотрят

Полнота покрытия распознавания

- 84% на чистых фрагментах
- 46% на шумных фрагментов
- 17% на фоновой музыке

Выводы

- Оценка качества ML не так сильно отличается от оценки качества кода
- Вкладывайтесь в тестовое окружение
- Вкладывайтесь в ручное тестирование
- Придумывайте метрики исходя из задачи

Выводы

- Оценка качества ML не так сильно отличается от оценки качества кода
- Вкладывайтесь в тестовое окружение
- Вкладывайтесь в ручное тестирование
- Придумывайте метрики исходя из задачи

Спасибо!