

Санкт-Петербургский государственный университет

Кафедра технологии программирования

Пичушкин Константин Андреевич

Гибридная система классификации музыкальных треков

Бакалаврская работа

Научный руководитель:
старший преподаватель Мишенин Алексей Николаевич

Рецензент:
Юршина Анастасия Александровна

Санкт-Петербург
2017

SAINT-PETERSBURG STATE UNIVERSITY

Chair of Programming Technology

Pichushkin Konstantin Adreevich

A hybrid system for music classification

Graduation Thesis

Scientific supervisor:
Senior Lecturer Mishenin Alexey Nikolaevich

Reviewer:
Yurshina Anastasia Alexandrovna

Saint-Petersburg
2017

Оглавление

| | |
|--|-----------|
| Введение | 4 |
| Постановка Задачи | 5 |
| 1. Обработка и хранение характеристик | 6 |
| 1.1. Выбор треков | 6 |
| 1.2. Выделение характеристик | 7 |
| 1.2.1. Музыкальные теги | 7 |
| 1.2.2. Аккорды | 7 |
| 1.2.3. Остальные признаки | 8 |
| 2. Выбор алгоритмов обучения рек. системы | 9 |
| 2.1. Кластеризация | 9 |
| 2.1.1. Walktrap | 9 |
| 2.2. Персонализированный pagerank | 11 |
| 3. Тестирование и эксперимент | 12 |
| 3.1. Результаты анализа данных | 12 |
| 3.2. Эксперименты и результат | 12 |
| Заключение | 13 |
| Список литературы | 14 |

Введение

Пожалуй, каждый среднестатистический пользователь интернета встречался с рекомендательными системами. Данный подход используется в рекламе, интернет магазинах, различных интернет проектах, посвященных литературе, музыке и кинематографу. Рекомендательная система – это программа, которая пытается предсказать, какие объекты будут интересны пользователю.

Наиболее распространенные подходы для создания рекомендательных систем:

–коллаборативная фильтрация использует данные о поведении пользователя в прошлом для создания рекомендаций. Помимо данных о конкретном пользователе, анализируются действия других пользователей со схожими характеристиками.

–фильтрация на основе содержания(контенте) использует данные об объектах для построения рекомендаций. Анализ фильтрации на основе содержания музыкальных композиций являлся темой моей преддипломной практики.

–в гибридных системах используются одновременно оба вида фильтрации, как правило, с помощью таких систем достигается более качественный подбор объектов.

В данной работе будет описан процесс создания гибридной рекомендательной системы музыкальных треков. В данном контексте под гибридностью подразумевается возможность добавления в систему признаков, полученных как и с помощью коллаборативной фильтрации, так и с помощью фильтрации на основе содержания.

Постановка Задачи

Целью данной работы является реализация Гибридной музыкальной рекомендательной системы, основанной на коллаборативной фильтрации и фильтрации на основе содержания. Помимо этого система должна поддерживать добавление удаление и изменение признаков.

Данную задачу можно разбить на следующие подзадачи:

- Выделение музыкальных характеристик
- Получение данных о поведении пользователей
- Сбор и анализ выделенных данных
- Построение гибридной реккомендательной системы

1. Обработка и хранение характеристик

1.1. Выбор треков

Для решения поставленной задачи необходимо выделить основные популярные музыкальные жанры. Сбор данных с различных сайтов (wikipedia, last.fm.com, pitchfork.com, spotify.com) позволил выделить 60 наиболее часто прослушиваемых жанров:

Alternative Rock, Experimental Rock, Goth Rock, Grunge, Hard Rock, Indie Rock, New Wave, Progressive Rock, Punk, Shoegaze, Blues, Classical, Minimalism, Country, Breakbeat, Dubstep, Garage, Hard Dance, House, Drum'n'bass, Techno, Trance, Easy Listening, Lounge, Swing, Electronic, Ambient, Downtempo, IDM, Rap, Indie, Pop, Industrial, Ska, Jazz, Latino, New Age, Britpop, PopRock, Soft Rock, pop, R&B, Disco, Funk, Soul, Reggae, Dub, Glam Rock, Metal, Psychedelic, Rock & Roll, Folk, Cloud rap, Math rock, Post rock, Stoner, Grime, Jungle, Emo.

Last.fm — рекомендательный сервис для любителей музыки, предоставляющий пользователям возможность хранения статистики прослушанных музыкальных композиций, на основании которой в последующем формируются рекомендации. Также это крупнейшая база данных музыкальных произведений и исполнителей. Каждая музыкальная композиция характеризуется тегами, которые ”прикрепляют” к ней пользователи (подробнее о тегах речь пойдёт ниже). На сайте last.fm.com предоставлена возможность поиска музыкальных треков в рамках определённого жанра и выделения самых популярных композиций среди результатов запроса.

Далее, были выделены самые часто прослушиваемые композиции в каждом из выбранных жанров с помощью средств сайта last.fm.com. Таким образом, было выделено около 10 тысяч музыкальных композиций, которые будут использоваться в рекомендательной системе.

1.2. Выделение характеристик

1.2.1. Музыкальные теги

Ключевая разница между музыкальными жанрами и тегами в том, что теги, ”прикреплённые” к композиции формируют сами пользователи сети. Помимо принадлежности к определённому жанру, могут отмечаться субъективные характеристики музыкального трека (настроение, танцевальность и т.д.), также возможна классификация, основанная на приемлемости прослушивания аудио композиции во время определённого занятия (спорт, учёба, йога и т.д.), либо тексте песни. В целом, пользователь никак не ограничен в присвоении тегов к композиции, что позволяет дополнять жанровую классификацию.

Основываясь на данных с сайта lasfm.com было получено по 10 самых часто используемых тегов для каждой выделенных ранее композиций.

1.2.2. Аккорды

Одно из допустимых определений музыки – набор звуков, организованных главным образом во времени и по высоте. Нота – графическое обозначение звука, один из основных символов современной музыкальной нотации. Аккорд – сочетание нескольких музыкальных звуков различной высоты, звучащих одновременно. Существуют музыкальные композиции, где в каждый момент времени звучит не больше одной ноты (пение одного человека без инструментального сопровождения), но, как правило, источников звуков несколько, причём высота звуков различна.

Таким образом, каждую музыкальную композицию длительности T можно представить в виде функции $\Phi(t)$ при Ch , где (t при $[0..T]$), Ch = множество аккордов. Такое представление было получено для выбранных композиций посредством библиотеки `rumir`.

1.2.3. Остальные признаки

Spotify — первый стриминговый сервис, позволяющий бесплатно и легально прослушивать музыкальные композиции, не скачивая их на компьютер. Реализован поиск артистов, музыкальных плейлистов и композиций. Кроме того, в Spotify присутствуют музыкальные рекомендации, основанные посредством Echonest. Echo Nest - это музыкальная платформа для разработчиков и медиакомпаний, основанная на исследовании аудио и текстового содержания музыкальных композиций. Предоставлена возможность использования музыкальных рекомендаций, создания списков воспроизведения и различные средства анализа для потребителей и разработчиков. В 2016 году произошла миграция многих API-интерфейсов (интерфейсов программирования приложений) Echo Nest в API Spotify. С помощью средств данной платформы были выделены следующие музыкальные характеристики (подробное описание признаков будет предоставлено ниже):

Tempo, Key, Duration, Time signature, Beat structure, Downbeats, Segments, Timbre, Pitch, Loudness, Energy, Danceability,

2. Выбор алгоритмов обучения рек. системы

2.1. Кластеризация

Задачей кластеризации объектов является разбиение множества элементов на плотные, удаленные друг от друга подмножества. Алгоритмом кластеризации называют функцию, которая каждому элементу ставит в соответствие метку определённого кластера. В зависимости от задачи, кластеризация может преследовать различные цели:

1. Выделение типичных представителей каждого кластера для возможной последующей классификации.
2. Изучение структуры множества объектов.
3. Обнаружение нетипичных элементов в множестве. В нашем случае кластеризация ведётся с целью группирования подобных музыкальных композиций.

2.1.1. Walktrap

Граф G опишем с помощью матрицы связности.

Рассмотрим процесс случайного блуждания на графе G . В каждый момент процесса блуждающий объект находится в вершине и перемещается в вершину, выбранную случайным равновероятным образом среди её соседей. Последовательность посещенных вершин – это цепь маркова, состояния которой – это вершины графа. На каждом шаге вероятность перехода из вершины i в j равна $P_{i,j} = \frac{A_{i,j}}{d_i}$. Таким образом определяется матрица перехода P процесса случайного блуждания. Вероятность перехода из вершины i в j за t ходов обозначим как P_{ij}^t . $D = \text{diag}(d_k), k = 1..n, d_k$ - степени вершин.

Для того, чтобы сгруппировать вершины в кластеры, введём метрику на вершинах. Интуитивно метрика отражает близость вершин.

Введём метрику следующим образом:

$$r_{i,j} = \sqrt{\sum_{k=1}^n \frac{(P_{ik}^t - P_{jk}^t)^2}{d_k}} = \|D^{-\frac{1}{2}}P_{i\bullet}^t - D^{-\frac{1}{2}}P_{j\bullet}^t\|$$

Где $\|\cdot\|$ - евклидова норма, $P_{i\bullet}^t$ - вектор-столбец у которого в j позиции стоит P_{ij}^t . Данная метрица обобщается на кластеры обозначим вероятность добраться из кластера C в вершину j за t шагов за P_{Cj}^t , которая равна:

$$P_{Cj}^t = \frac{1}{|C|} \sum_{i \in C} P_{ij}^t$$

Расстояние между кластерами определяется как:

$$r_{C_1 C_2} = \|D^{-\frac{1}{2}}P_{C_1\bullet}^t - D^{-\frac{1}{2}}P_{C_2\bullet}^t\| = \sqrt{\sum_{k=1}^n \frac{(P_{C_1k}^t - P_{C_2k}^t)^2}{d_k}}$$

Теперь, когда задана метрика, задача выделения кластеров сведена к задаче кластеризации вершин. Далее приводится алгоритм, основой которого является метод Варда.

1. Каждая вершина принадлежит своему кластеру. $P_1 = \{v \mid v \in V\}$ - начальное разбиение.
2. Вычисление расстояния между смежными вершинами
3. На каждом k -ом шаге.
 - (a) Выбрать два кластера C_1 и C_2 из P_k исходя из определённого критерия
 - (b) Образовать один кластер $C_3 = C_2 \cup C_1$
 - (c) Вычислить расстояние между кластерами

Алгоритм завершется на $n - 1$ шаге. Описание алгоритма:

Этап (a): Рассматриваются инцидентные кластеры. Также данные кластеры должны минимизировать изменение среднего квадратов расстояния между каждой вершиной и их сообществом при объединении этих

сообществ (метод Варда):

$$\sigma_k = \frac{1}{n} \sum_{C \in P_k} \sum_{i \in C} r_{iC}^2 \rightarrow \min$$

Задача может быть сведена к

$$\Delta\sigma(C_1, C_2) = \frac{1}{n} \left(\sum_{i \in C_3} r_{iC_3}^2 - \sum_{i \in C_1} r_{iC_1}^2 - \sum_{i \in C_2} r_{iC_2}^2 \right) \rightarrow \min\{C_1, C_2\}$$

Таким образом мы используем жадный алгоритм для решения NP -трудной задачи: "KMedian clustering problem". Асимптотика данного метода $O(n^2 \log n)$.

2.2. Персонализированный pagerank

Этот алгоритм представляет из себя модификацию процесса случайного блуждания на графе [1]. Для реализации алгоритма необходимо выбрать m вершин графа A . Также определим число a – вероятность начать обход заново. На каждом шаге алгоритма с вероятностью $(1-a)$ блуждающий объект равновероятным случайным образом возвращается в одну из выбранных m вершин, с вероятностью a происходит переход в соседнюю вершину (в этом случае вероятность перехода в вершину j равняется $P_{i,j}$).

Таким образом, после окончания процесса блуждания, можно выделить вершины, которые посещались наибольшее количество раз. Эти вершины будут являться рекомендацией для пользователя, отметившего m вершин.

3. Тестирование и эксперимент

3.1. Результаты анализа данных

В целом, все характеристики показали намного менее точные результаты работы по сравнению с тегами. Ниже приведены результаты среднего квадратического отклонения характеристик. Первое число – среднее значение в сформированных кластерах, второе значение – среднее значение в случайных кластерах. `liveness(0.132323, 0.152323)`
`speechiness(0.042671243989, 0.063199034586)`, `acousticness(0.1839, 0.19080584678)`
`instrumentalness(0.1673102, 0.187527)`. Можно заметить, что разница совсем невелика.

Такие же результаты показали характеристики после введения метрики, в среднем теги показывали точность в 400 раз лучшую, чем любые другие выделенные показатели

3.2. Эксперименты и результат

Начальная точность алгоритма, без введения метрик 22.232 процента, после введения 27.332. При этом добавление признаков не дало сильного улучшения работы, но этот результат был ожидаем после анализа. Система была основана на персонализированном pagerank.

Заключение

В данной работе была реализована рекомендательная система аудио треков, поддерживающая изменение, добавление и удаление признаков. В ходе работы были проанализированы различные характеристики. Основываясь на анализе, можно сделать вывод, что музыкальные теги – один из основных признаков, приемлемых для создания рекомендательной системы. Одним из возможных путей развития данной системы является добавление большего количества характеристик. [2]

Список литературы

- [1] Edge-weighted personalized pagerank: Breaking a decade-old performance barrier / Wenlei Xie, David Bindel, Alan Demers, Johannes Gehrke // Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining / ACM. — 2015. — P. 1325–1334.
- [2] Wikipedia. Music Genome Project — Wikipedia, The Free Encyclopedia. — 2017. — [Online; accessed 25-May-2017]. URL: https://en.wikipedia.org/w/index.php?title=Music_Genome_Project&oldid=779229313.