

Методы рекомендации музыки для терапии на основе данных электроэнцефалографии

Полевой А.В.

Московский государственный университет имени
М.В.Ломоносова
Москва, Россия
polevoianton@bk.ru

Корухова Ю.С.

Московский государственный университет имени
М.В.Ломоносова
Москва, Россия
yulia@cs.msu.ru

Аннотация

В современном мире цифровые технологии используются в различных предметных областях, не остается в стороне и медицина. С появлением больших наборов оцифрованных данных возникают задачи их автоматической обработки для принятия решений. Одним из актуальных направлений является музыкальная терапия. Новизна данной работы состоит в повышении качества составления рекомендаций для музыкальной терапии на основе ЭЭГ. В работе предложен подход к решению проблемы холодного старта, основанный на обученной модели по набору Musin-G. Таким образом, в рамках работы предложен подход для построения рекомендательной системы с использованием модели данных ЭЭГ для задач музыкальной терапии. Результаты экспериментов показали, что наилучшая точность 62.5% достигается у моделей градиентного бустинга.

1 Введение

Электроэнцефалография - это раздел электрофизиологии, изучающий закономерности суммарной электрической активности мозга, отводимой с поверхности кожи волосистой части головы. Она дает возможность качественного и количественного анализа функционального состояния головного мозга и его реакций при действии раздражителей. В задачах рекомендации музыки в медицине ставится задача подобрать набор музыкальных композиций, наилучшим образом подходящих для терапии.

Основной целью данного исследования является повышение качества существующих рекомендательных систем за счёт анализа данных электроэнцефалограмм.

Статья организована следующим образом. В главе 2 кратко представлены используемые наборы данных. Далее, в главе 3 опишем используемые метрики и методы. Затем предложены варианты развития для некоторых методов. В заключительных главах опишем эксперименты и проанализируем полученные результаты.

2 Обзор данных

Известные системы составления рекомендаций по данным ЭЭГ оценивались на небольших медицинских данных или закрытых данных, что несёт определенные ограничения в интерпретируемости результатов. Новизна в данной работе состоит в проведенном авторами систематическом сравнении

точности широкого набора алгоритмов на различных наборах данных академической направленности. Остановимся подробнее на каждом наборе:

Musin-G. Набор данных [6] содержит результаты замеров электроэнцефалографии 20 индийских участников на 12 песен различных жанров (от индийской классики до готик-рока). В каждой сессии песня обозначена номером.

Для проведения эксперимента участникам предлагалось закрыть глаза, о чем свидетельствовал одиночный звуковой сигнал, и песня предъявлялась им через динамики. После прослушивания каждой песни раздавался двойной звуковой сигнал, просивший участников открыть глаза и оценить степень знакомства и удовольствия от песни. Ответы давались по шкале от 1 до 5, где 1 означало, что песня наиболее знакома или наиболее приятна, а 5 - наименее знакома или наименее приятна.

DEAP. Мультимодальный набор данных [3] для анализа аффективных состояний человека. ЭЭГ и периферические физиологические сигналы 32 участников были записаны во время просмотра 40 одномоментных отрывков музыкальных видеоклипов. Участники оценивали каждое видео по уровням возбуждения, валентности, симпатии/антипатии, доминирования и знакомости. Для 22 из 32 участников также было записано фронтальное видео лица.

DECAF. Мультимодальный набор данных DECAF [4] содержит сигналы мозга, полученные с помощью датчика магнитоэнцефалограммы (МЭГ), который не требует физического контакта с пользователем. В нём отражены эмоциональные реакции 30 участников на 40 одномоментных музыкальных видеофрагментов, что позволяет проводить сравнение между ЭЭГ и МЭГ. В дополнение к данным МЭГ, DECAF включает в себя синхронно записанные видео лица, горизонтальную электроокулограмму (hEOG), электрокардиограмму (ECG) и периферические физиологические реакции трапецевидной мышцы-электромиограммы (tEMG).

NMED-M. Этот набор данных [2] [5] содержит результаты замеров ЭЭГ и поведенческие данные, собранные во время естественного прослушивания музыки. Электроэнцефалограмма была записана у 30 взрослых участников, каждый из которых слышал набор из пяти музыкальных композиций. Участники оценивали, насколько каждый стимул был приятным и музыкальным.

В рамках текущей работы нами рассмотрен набор Musin-G. Данный выбор был аргументирован разнообразием представленных музыкальных жанров, а также актуальностью набора. Набор был выложен в открытый доступ в 2021 году и в рамках

Название	DEAP	DECAF	NMEDM	Musin-G
Размер	32 участника, 40 видео (1 мин. каждое) 2.7 ГБ (обработанный)	30 участников	30 участников, 5 композиций 34 ГБ	20 участников, 12 композиций 10,1 ГБ
Стимул	• Музыкальные видео	• Музыкальные видео	• Музыка	• Музыка
Технология	• ЭЭГ	• МЭГ	• ЭЭГ	• ЭЭГ
Авторы	• Sander Koelstra, Christian Muehl et al.	• Mojtaba Khomami Abadi, Ramanathan Subramanian et al.	• Dauer, Tysen, Nguyen et al.	• Krishna Prasad Miyapuram, Pankaj Pandey et al.

Рис. 1: Сравнительная таблица наборов данных

обзора не было найдено открытых исследований по анализу этого набора.

3 Методы

Проведём обзор популярных подходов в литературе, для обоснования дальнейшего развития методов коллаборативной фильтрации и решению проблемы холодного старта.

Опираясь на работы, которые систематизируют существующие подходы в рекомендательных системах, можно поделить существующие методы на две группы.

- **Коллаборативная фильтрация**

Алгоритмы коллаборативной фильтрации используют информацию о поведении пользователей в прошлом – например, информацию о прослушанных ранее композициях. В этом случае учитываются неявные характеристики, которые было бы невозможно учесть при создании профиля. Однако проблемой коллаборативной фильтрации является так называемый «холодный старт»: отсутствие данных о новых пользователях или объектах. Коллаборативная фильтрация применяется во многих проектах, таких как Last.fm, Amazon, Wikimart.

- **Фильтрация содержимого**

Методы фильтрации содержимого основаны на использовании дополнительной информации об объектах. Рекомендации составляются путем сопоставления этой информации с профилями пользователей. В случае музыкальной рекомендательной системы дополнительной информацией о композициях может служить жанр, название исполнителя, год записи и т.д., тогда как в профиле пользователя могут содержаться указанные музыкальные вкусы и любимые композиции. Основной проблемой является то, что пользователям для получения рекомендаций приходится много рассказывать о себе, тогда как методы коллаборативной фильтрации работают автономно.

По итогам обзора было принято решение развивать идеи методов, связанных с коллаборативной фильтрацией и решать проблемы холодного старта при помощи модели для оценки

Подход	Преимущества	Недостатки
Неперсонализированный подход	Простота реализации.	Не учитывает предпочтения отдельного пользователя.
Фильтрация на основе содержания	Учет необычных вкусов пользователей. Не нужна большая группа пользователей. Возможность рекомендовать неочевидные элементы.	Проблема нового пользователя. Отсутствие разнообразия в рекомендациях.
Коллаборативная фильтрация	Универсальность. Разнообразные рекомендации. Не нужно много информации о пользователях и элементах.	Проблема холодного старта. Проблема первого оценщика. Разреженность. Смещение популярности.

Рис. 2: Сравнение алгоритмов для рекомендательных систем

вовлеченности слушателя по данным ЭЭГ. Поскольку в задачах музыкальной терапии пользователю предпочтительнее подбирать объекты согласно его состоянию, а не по желанию.

Метрики.

4 Развитие методов

Решение проблемы холодного старта.

Нами было предложено использовать модель, определяющую степень соответствия композиции к ЭЭГ пользователя, для пополнения начального тренировочного набора в рекомендательной системе.

Подробная схема предлагаемого подхода представлена на рис.3. Предположим, что у нас есть модель, которая по данным ЭЭГ, определяет уровень приятных и знакомых композиций. Тогда для рекомендации человеку какой-то композиции из известных, для начала найдем близких к нему по ЭЭГ, если такие окажутся. Тогда рекомендацией для него будет подмножество композиций подходящих для данных пользователей. Если при этом человек первый в нашей работе, или для него не нашлось похожих пользователей, то тогда мы предлагаем послушать ему случайную композицию и снимаем данные ЭЭГ. На полученных данных делаем прогноз. Если модель уверена в соответствии данной композиции для человека, то мы добавляем в нашу базу данное наблюдение, в противном случае, выбор случайной композиции повторяется.

Модуль 1 ищет схожих к данному пользователю $f(x)$ для всех доступных $g(x)$ по ЭЭГ при помощи скалярного произведения:

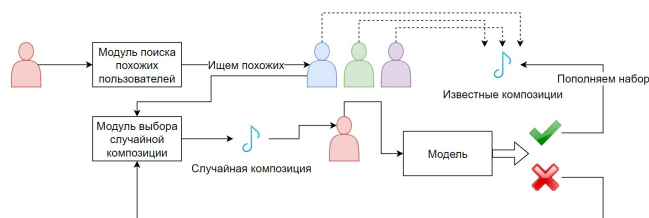
$$\int_0^T f(x) * g(x) dx$$

Модуль 2 осуществляет выбор композиции из доступного множества композиций.

5 Эксперименты

В экспериментах использовались наборы данных, описанные в разделе 2. Для реализации методов в качестве классификатора был использован CatBoost [1].

Проводились эксперименты в различных конфигурациях: Musin-G (903 признака), Musin-G (516 признаков), Musin-G (177



Lomas. 2021. "Music Listening- Genre EEG dataset (MUSIN-G)". <https://doi.org/10.18112/openneuro.ds003774.v1.0.0>

Рис. 3: Схема работы системы для составления рекомендаций

отобранных признаков). Далее для всех трёх конфигураций выполнялись следующие действия:

- Наборы делились на тренировочную (0.7) и тестовую (0.3).
- Параметры подбирались кросс валидацией на тренировочной выборке.
- Для подходов использовался CatboostClassifier [1]
- Качество моделей оценивалось по метрике Accuracy, результаты усреднялись по 100 случайным подвыборкам из тестового набора.
- Качество моделей оценивалось единообразно с оптимизацией на функцию потерь многоклассовой классификации.

Проблема с переобучением градиентных бустингов решалась при помощи настройки размера шага и числа деревьев, а также бэггинга (обучение базовых алгоритмов по случайным подвыборкам). А также регуляризации за счет константы `l2_leaf_reg`.

На основе экспериментов был сделан вывод, что необходимо использовать наборы с меньшим числом признаков и делать трансформацию целевой переменной.

6 Заключение

Таким образом, в рамках работы предложен подход для построения рекомендательной системы с использованием модели данных ЭЭГ для задач музыкальной терапии. Результаты экспериментов показали, что наилучшая точность 62.5% достигается у моделей градиентного бустинга.

Список литературы

- [1] [n.d.]. CatBoost Classifier, <https://github.com/catboost/catboost>.
- [2] [n.d.]. Dauer, Tysen, et al. "Naturalistic music EEG dataset—Minimalism (NMED-M)." (2021).
- [3] [n.d.]. "DEAP: A Database for Emotion Analysis using Physiological Signals (PDF) S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, 2012.
- [4] Mojtaba Khomami Abadi, Ramanathan Subramanian, Seyed Mostafa Kia, Paolo Avesani, Ioannis Patras, and Nicu Sebe. 2015. DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses. *IEEE Transactions on Affective Computing* 6, 3 (2015), 209–222. <https://doi.org/10.1109/TAFFC.2015.2392932>
- [5] Tysen Dauer, Duc T. Nguyen, Nick Gang, Jacek P. Dmochowski, Jonathan Berger, and Blair Kaneshiro. 2021. Inter-Subject Correlation during New Music Listening: A Study of Electrophysiological and Behavioral Responses to Steve Reich's Piano Phase. *bioRxiv* (2021). <https://doi.org/10.1101/2021.04.27.441708> arXiv:<https://www.biorxiv.org/content/early/2021/04/28/2021.04.27.441708.full.pdf>
- [6] Krishna Prasad Miyapuram, Pankaj Pandey, Nashra Ahmad, Bharatesh R Shiraguppi, Esha Sharma, Prashant Lawhatre, Dhananjay Sonawane, and Derek