

**ОБЗОР ПОДХОДОВ ДЛЯ РЕШЕНИЯ
ЗАДАЧИ АННОТИРОВАНИЯ МУЗЫКИ**

Иванченко А.А. (ИТМО)

**Научный руководитель – доцент Бойцев А.А.
(ИТМО)**

Введение. Музыка – это форма общения, которая может отражать человеческие эмоции, географическое происхождение, духовные основы и другие аспекты человечества. Иногда для музыкальных произведений создают описания или аннотации. Подобно другим формам художественного самовыражения, анализ музыки можно рассматривать как поиск смысла, который композиторы заложили в своих произведениях. Аннотированная музыка предоставляет ценные данные для обучения и оценки алгоритмических методов поиска музыкальной информации (MIR) и компьютерного анализа музыки (CMA), а также она может быть использована для создания инновационных решений, обеспечивающих доступ к музыкальному контенту для людей со слабым слухом.

Основная часть. На данный момент существуют большие открытые наборы данных с определенными тегами для аудио, к примеру: жанр, инструмент, темп и т.п. Поэтому задача аннотирования музыки на сегодняшний день представляет из себя в основном мультиклассовую классификацию. Наиболее распространёнными методами решения данной задачи являются: метод опорных векторов (SVM) [1] и сверточные нейронные сети [3].

В первом случае для каждой дорожки в аудиокolleкции вычисляется вектор признаков на основе аудиоконтента. Поскольку каждая дорожка может быть аннотирована несколькими тегами, вектор признаков подается в мультиклассовую звуковую SVM. После обработки всех дорожек данные передаются в линейную SVM для обучения, так вычисляется вектор сходства тегов [1].

Второй метод основывается на классификации изображений. То есть для решения берутся классические сверточные нейронные сети, которые адаптируются для аудиоданных. На вход моделям подаются спектрограммы аудио, по которым происходит мультиклассовая классификация [2].

Помимо этого, появляются новые наборы данных, в которых появляется полноценное описание музыки в нескольких предложениях, что помогает лучше понять её [3].

Выводы. Таким образом были рассмотрены основные подходы для решения задачи аннотирования музыки и представлены различные наборы данных. Так как музыкальное произведение несет в себе определенный скрытый смысл, аннотирование музыки не может быть полностью сведено к решению задачи классификации.

Список использованных источников:

1. Ness S., Theocharis A., Tzanetakis G. Improving Automatic Music Tag Annotation Using Stacked Generalization of Probabilistic SVM outputs // Proceedings of the 17th ACM international conference on Multimedia – 2009. – P. 705-708.
2. Damböck, M. A Comparison of Audio Preprocessing Methods for Music Autotagging using CNN-architectures // Diploma Thesis, Technische Universität Wien. reposiTUm. – 2022 – 100 p. – URL: <https://doi.org/10.34726/hss.2022.89400>
3. Ilaria M., Benno W., Seunghoon D. The Song Describer Dataset: A Corpus of Audio Captions for Music-and-Language Evaluation // eprint arXiv: – 2023 – 13 p. – URL: <https://arxiv.org/pdf/2311.10057.pdf>

Иванченко А.А. (автор)

Подпись

Бойцев А.А. (научный руководитель)

Подпись