

0.1. Поляков С.А., Казаков Г., Иванов К.О. Разработка модели машинного обучения для автоматической генерации гитарных табулатур

В данной исследовательской работе в области глубинного обучения в обработке звука, представлена разработка модели машинного обучения, целью которой является автоматическая генерация гитарных табулатур на основе аудиофайлов. Этот проект имеет важное практическое применение в мире музыки, предоставляя музыкантам и композиторам инструмент для более эффективной работы с музыкальными данными. Разработанная модель может быть использована для автоматической транскрипции музыки, что экономит время и усилия музыкантов при создании нотных записей. Кроме того, она может быть востребована для создания гитарных аранжировок, открывая новые возможности в области музыкального творчества.

Для обучения модели были использованы следующие датасеты:

MAESTRO: Датасет, содержащий классическую музыку, который предоставляет разнообразие стилей и инструментов.

MusicNet Dataset: Этот датасет включает разнообразные музыкальные произведения и предоставляет ценную информацию для обучения модели.

GuitarSet: Специализированный датасет для гитаристов, содержащий записи гитарных звуков и метаданные.

Самостоятельно Синтезированные Датасеты: Дополнительно были созданы датасеты, включающие аудиофайлы и соответствующие midi файлы.

Также для обучения и достижения хороших результатов были использованы библиотеки: Librosa, PyTorch, Music21, Matplotlib, Numpy, Pandas, а другие стандартные библиотеки языка программирования Python.

Вместо обучения на целых музыкальных файлах было разработано решение, при котором аудиофайл разделяется на паттерны, такие как вокал, бас, барабаны и аккомпанемент. Для каждого паттерна генерируется соответствующий MIDI файл. Это позволяет более точно извлекать информацию о мелодии и аккордах из аудиофайла. С помощью специального алгоритма

В процессе разработки модели возникли следующие проблемы:

1) Наличие шумов и дополнительных музыкальных эффектов в аудиофайлах оказывает существенное влияние на качество предсказаний.

2) Сложность разделения аудио-файла на паттерны. Для достижения максимальной точности в процессе реализации нашей модели мы применили несколько важных методов. В частности, мы использовали преобразование Фурье, которое позволило улучшить качество обработки аудиоданных. Кроме то-

го, мы предложили использовать не только мел-спектрограммы, но и хроматограммы для обучения модели. Это решение существенно повысило эффективность и точность нашей модели в анализе звуковых данных.

В целом наша модель показала неплохие результаты по придуманным нами метрикам. В дальнейшем планируется улучшить качество модели за счет увеличения количества тренировочных данных.

Литература: <https://paperswithcode.com/task/music-transcription> - 33 научных статьи посвященных музыкальной транскрипции

<https://pytorch.org/> - Сайт открытой библиотеки Pytorch

<https://librosa.org/doc/latest/index.html> - Документация библиотеки librosa