

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.032.26:78.073

Швед
Елизавета Ивановна

Программное средство для анализа музыкальных предпочтений пользователя
с использованием нейросетевых алгоритмов

АВТОРЕФЕРАТ
на соискание степени магистра
по специальности 1-40 80 01 «Компьютерная инженерия (встраиваемые
системы)»

(подпись магистранта)

Научный руководитель
Азаров Илья Сергеевич

(фамилия, имя, отчество)

Доктор технических наук, доцент

(ученая степень, ученое звание)

(подпись научного руководителя)

Минск 2024

ВВЕДЕНИЕ

Музыка – это звук, организованный специальным образом для выражения различных идей и эмоций. Столь большое внимание к музыке во многом объясняется ее цифровизацией и распространением портативных аудиовоспроизводящих устройств, что существенно упрощает доступ к музыкальному материалу. Цифровизация музыки предоставила практически неограниченные возможности для реализации «свободы творчества» и огромный выбор, а также возможность создания подборок, основываясь на уже понравившихся произведениях. У аудиоданных есть ряд свойств, которые, в некотором роде, делают их похожими на то, что обычно изучается в рамках глубокого обучения: последовательный характер музыки напоминает NLP; композиция может иметь несколько «каналов» звука с точки зрения тонов и инструментов, которые напоминают изображения.

Нейронные сети представляют собой набор алгоритмов, которые пытаются распознать отношения и связи между входными данными, имитируя работу человеческого мозга. Они интерпретируют данные посредством своего рода машинного восприятия, маркируя или группируя исходные данные. Образцы, которые они распознают, являются числовыми, содержащимися в векторах, в которые должны быть переведены все данные реального мира, будь то изображения, звук, текст или временные ряды. Таким образом, определение жанра музыки в общем смысле зависит главным образом от двух факторов. Первым из них является представление данных, то есть кодирование входных данных в форму, наиболее удобную и эффективную для обучения, а также интерпретация полученных результатов. Вторым фактором является выбор алгоритма нейросетевого моделирования, которые может эффективно работать с предоставленной информацией, и таким образом обеспечит выходными данными, определяющими музыкальные предпочтения пользователя.

В данной работе необходимо изучить, реализовать и проанализировать различные представления данных, а также алгоритмы нейросетевого моделирования, которые позволят решить поставленную задачу.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цели и задачи исследования

Целью диссертации является исследование и анализ применения нейросетевых моделей и алгоритмов, которые позволяют по выделенным в рамках исследования параметрам анализировать стиль музыкальных композиций.

Для достижения поставленной цели необходимо решить следующие задачи:

- провести сравнительный анализ нейросетевых моделей для прогнозирования результатов обучения;
- реализовать модели для рекомендации музыкальных композиций;
- провести оценку эффективности использования нейросетевых моделей для рекомендации музыкальных композиций.

Область исследования

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) ОСВО 1-40 80 01-2019 специальности 1-40 80 01 Компьютерная инженерия (профилизация Встраиваемые системы).

Теоретическая и методологическая основа исследования

В основу диссертации легли работы зарубежных ученых в области методов машинного обучения, а также анализа образовательных данных.

Информационная база исследования сформирована на основе литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Научная новизна, теоретическая и практическая значимость

Научная новизна и значимость полученных результатов работы заключается в разработке моделей машинного обучения для рекомендации музыкальных композиций.

Теоретическая значимость работы заключается в обзоре современных методов машинного обучения для работы музыкой, их математических моделей и базовых алгоритмов. Оценки наиболее подходящих для поставленных задач моделей.

Практическая значимость диссертации состоит в разработке нейросетевой модели для анализа и составления рекомендаций музыкальных композиций на основе предпочтений и поведения пользователей.

Личный вклад соискателя

Результаты, полученные в результате написания диссертации, получены соискателем лично. Вклад научного руководителя Азарова И.С., заключался в формулировке целей и задач исследования и консультаций по теме.

Публикации

По теме диссертации опубликовано в 2 печатные работы, все из которых являются статьями в сборниках материалов конференций БГУИР.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, пяти глав, заключения, списка использованных источников, списка публикаций автора и приложений.

В первой главе представлен обзор предметной области исследования, область применения, обзор моделей нейросетевых алгоритмов, рассмотрены возможные используемые технологии.

Во второй главе описывается методология, используемая для достижения целей данного исследования, а также проводится анализ и оценка исходных данных .

Третья глава включает рассмотрение нескольких алгоритмов музыкальных рекомендаций, основанных на свёрточных нейронных сетях и их сравнение.

Четвертая глава описывает модель рекомендаций, основанная на свёрточной нейронной сети, обучение и экспериментальную среду.

В пятой главе описано объединение экспериментальной модели системы пользовательских предпочтений и результатов ее анализа.

Общий объем диссертационной работы составляет 72 страницы. Из них 52 страниц основного текста, 26 иллюстраций на 19 страницах, библиографический список из 20 наименований на 2 страницах, список собственных публикаций соискателя из 2 наименований на 1 странице, 2 приложения на 16 страницах.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

По результатам обзора, описанного в первой главе работы, были описаны виды нейронных сетей и их преимущества. Было определено, что целевая задача составления представления о музыкальных предпочтениях слушателей по прослушанным аудиозаписям сводится к задаче выделения абстрактных характеристик у композиций, таких как жанр музыки. Данная задача является классическим случаем задачи классификации, которую можно решить с использованием нейросетевых алгоритмов. Было исследовано, что сверточные нейронные сети хорошо приспособлены к решению задач данного типа, а именно, составление характеристических карт абстрактных признаков по входным данным. Также было определено, что рекуррентные нейронные сети и их усовершенствованные виды, такие как LSTM, способны анализировать данные переменной длины, что подходит для решения задачи определения жанра.

Во второй главе объяснены все этапы, начиная с получения, подготовки, моделирования и оценки данных. В третьей главе описаны рекомендательные алгоритмы, расширяющие разработку на оригинальной основе, в основном они подразделяются на три категории: алгоритм рекомендации контента, алгоритм рекомендации контекста и алгоритм рекомендации глубокого обучения, в зависимости от объекта рекомендации. Контекстная информация может использоваться в качестве основы для коллаборативной фильтрации, чтобы оценить степень сходства между пользователями. Такие механизмы преобразования аудиозаписей, как конвертирование в мел-частотные кепстральные коэффициенты, способны существенно сжать анализируемые данные, убрать шумы и выделить те черты аудиозаписей, которые анализируются человеческим слухом в первую очередь. В четвертой главе описан процесс обучения сверточной нейронной сети, а также проведены сравнение и анализ функций активации.

В пятой главе проведено объединение экспериментальной модели системы предпочтений пользователя и результатов ее анализа. Показана взаимосвязь между пользователями и характеристиками распределения, проведена комплексная и мультикатегорийная оценка характеристик пользователей.

Алгоритм рекомендации сверточной нейронной сети на основе контента также использует обучение звуковых функций, но в отличие от этого метода заключается в том, что на основе контента требуется большое количество пользовательских оценок для прогнозирования популярности музыки, а затем рекомендации основаны на популярности.

Из результатов видно, что алгоритм рекомендаций, основанный на сверточной нейронной сети на основе контента, которая опирается на данные оценки пользователей, имеет низкую точность рекомендаций во время холодного старта, но точность рекомендаций начинает увеличиваться при добавлении данных оценки пользователей. Поскольку метод в этой работе не зависит от оценки пользователей, точность рекомендаций не сильно меняется. Таким образом, этот метод имеет определенное значение использования в отсутствие данных оценки пользователей для рекомендации.

В результате приведено несколько примеров рекомендаций пользователю по выбранной композиции с указанием схожих характеристик.

ЗАКЛЮЧЕНИЕ

В данной работе была разработана и протестирована рекомендательная система музыкальных треков, основанная на использовании сверточных нейронных сетей. Система позволяет генерировать персонализированные рекомендации для пользователей на основе их предыдущих музыкальных предпочтений.

Применение нейронных сетей продемонстрировало высокую эффективность в создании рекомендательных систем для музыкальных композиций. Модели, обученные на данных о прослушиваемой пользователями музыки, успешно предсказывали новые треки, которые могут понравиться конкретному человеку. Это открывает возможности для развития интеллектуальных систем, способных адаптировать музыкальные рекомендации в соответствии с текущим состоянием слушателя. Подобные рекомендательные системы могут помочь пользователям решить проблему однообразного прослушивания музыки, открывая им новые композиции, соответствующие их вкусам и настроению.

Диссертация состоит из введения, пяти глав, заключения, списков литературы и публикаций автора, а также приложений. Первая глава посвящена обзору предметной области, моделей нейросетевых алгоритмов и используемых технологий. Вторая глава описывает методологию исследования и анализ исходных данных. В третьей главе проведено сравнение нескольких алгоритмов музыкальных рекомендаций, основанных на сверточных нейронных сетях. Четвертая глава фокусируется на разработке модели рекомендаций с использованием сверточной нейронной сети, ее обучении и экспериментальной оценке. Пятая глава посвящена интеграции экспериментальной модели системы пользовательских предпочтений и анализу полученных результатов.

Данная работа демонстрирует перспективность применения нейронных сетей в создании интеллектуальных музыкальных рекомендательных систем, способных улучшить музыкальный опыт пользователей. Дальнейшие исследования в этом направлении могут способствовать большему развитию и усовершенствованию подобных систем.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

[1-А] 60-я Научная Конференция Аспирантов, Магистрантов и Студентов БГУИР – Швед Е.И., Тушинская Е.В., Рогов М.Г. – СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИНСТРУМЕНТОВ ДЛЯ СОЗДАНИЯ REALTIME-ГРАФИКИ НА ОСНОВЕ ОБРАБОТАННОГО АУДИОСИГНАЛА. – 22–26 апреля 2024 года, Минск, БГУИР. – Находится в процессе печати.

[2-А] 60-я Научная Конференция Аспирантов, Магистрантов и Студентов БГУИР – Швед Е.И., Сельханович М.А., Потейчук В.М., Рогов М.Г. – ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕЙ ДЛЯ АНАЛИЗА МУЗЫКАЛЬНЫХ ПРЕДПОЧТЕНИЙ. – 22–26 апреля 2024 года, Минск, БГУИР. – Находится в процессе печати.