

УДК 004.89

***ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ ОБРАБОТКИ  
И СОЗДАНИЯ МУЗЫКАЛЬНЫХ ПРОИЗВЕДЕНИЙ: МЕТОДЫ,  
ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ***

***Баврин А.Э.***

*студент*

*ФГБОУ ВО «Калужский государственный университет им. К. Э. Циолковского»  
Калуга, Россия*

***Салтыкова Н.В.***

*к.пед.н., доцент*

*ФГБОУ ВО «Калужский государственный университет им. К. Э. Циолковского»  
Калуга, Россия*

**Аннотация:** В статье рассматривается использование нейросетевых моделей для обработки и создания музыкальных произведений. Анализируются наиболее подходящие решения в применении ИИ-агентов в сфере музыки, как в контексте теоретического рассмотрения принципов и методов работы моделей с музыкальными композициями, так и в приведении конкретных примеров таких моделей и сравнения их преимуществ и ограничений по теме. Освещены такие новые явления в музыкальной сфере как ИИ-кавер и мэшап. Отражены дальнейшие перспективы развития и возникновения проблемных вопросов в применении нейросетей в работе с музыкой.

**Ключевые слова:** музыка, нейросеть, кавер, мэшап, ИИ

***USING NEURAL NETWORK MODELS FOR PROCESSING AND  
CREATING MUSICAL WORKS: METHODS, OPPORTUNITIES, AND  
PERSPECTIVES***

***Bavrin A.E.***

*Student*

*Kaluga State University named after K. E. Tsiolkovsky*

*Kaluga, Russia*

***Saltykova N.V.***

*PhD, Associate Professor*

*Kaluga State University named after K. E. Tsiolkovsky*

*Kaluga, Russia*

**Abstract:** The article discusses the use of neural network models for processing and creating musical compositions. It analyzes the most suitable solutions for using AI agents in the field of music, both in terms of theoretical analysis of the principles and methods of working with musical compositions, and in providing specific examples of such models and comparing their advantages and limitations. The article also covers new phenomena in the music industry, such as AI-generated covers and mashups. It also explores prospects and challenges in using neural networks for music processing.

**Key words:** music, neural network, cover, mashup, AI

Продолжительное время ИИ является одной из основных тем информационного поля ввиду больших перспектив, которые это направление открывает перед всем человечеством в самых разных сферах жизни. Нейросети привлекли в свою отрасль большие инвестиции в разработки, исследования и строительство крупных дата-центров, которые, в обозримом будущем, смогут существенно повысить вычислительные мощности, а следовательно, значительно расширить спектр направлений использования ИИ-агентов в жизни человека.

Как это уже случалось ранее, новые технологии смогут, по мере своего развития, внедриться не только в технические области, но и в более

«человеческие», связанные с эмоциями и абстракцией, например в искусство, в частности, в музыку. Аудиальное искусство тесно связано с чувственным восприятием человека, что делает тему сопряжения новейших технологий (нейросетей) и музыки крайне актуальной. В данной статье рассматриваются уже сформировавшиеся методы работы с музыкой при использовании нейросетевых инструментов, новые жанровые форматы, применяемые ИИ-сервисы, а также дальнейшие перспективы в области.

Цель данной статьи - рассмотреть основные музыкальные сервисы-редакторы, использующие нейросети и проанализировать возможности, предоставляемые начинающему пользователю нейросетевых инструментов.

Для реализации цели нами были поставлены следующие задачи:

1. Рассмотреть особенности работы нейросети с музыкальными произведениями;
2. Изучить основные музыкальные сервисы-редакторы, использующие нейросети;
3. Выбрать критерии для сравнения сервис-редакторов.

Для решения первой задачи остановимся подробнее на специфике работы нейросети с музыкальными произведениями. Мы рассматриваем случай, когда ИИ основывает свою работу по входному аудиофайлу. Процесс применения нейросети заключается в разделении дорожек, обработка отдельных фраз звуковыми эффектами, миксование отдельных отрывков или их дополнение, анализ входных файлов и генерация на их основе новых, модификация на основе полученных в ходе анализа композиции паттернов или уже имеющихся ранее на основе обучения нейросети в целом. [4]

Вопреки распространённому мнению нейросети не являются чем-то однородным. Существует множество видов нейросетей, каждый более подходящий для определённых типов задач. В рассмотрение темы, для работы с

музыкой лучше всего подходят рекуррентные нейросети (RNN), в частности, их тип – LSTM (long short-term memory), которые умеют анализировать последовательности данных, в которых важно, в каком порядке идут значения, при том сохраняя информацию о долгих зависимостях внутри анализируемых данных, что позволяет обрабатывать и достраивать новые данные в контексте исходных на долгих продолжительностях. Нейроны внутри сети реализуют между собой «память» о прежних состояниях сети после её изменений, что позволяет видеть долгие закономерности и достраивать их без существенных отклонений, что как нельзя лучше подходит для процесса обработки музыкальных произведений. [1]

Также, вместе с RNN применяются свёрточные нейросети (CNN). Они предназначены для анализа спектрограммы аудиофайла, что позволяет им довольно точно определять как базовые характеристики музыкальных фраз (bpm, ритм, тональность), так и эмоциональные (настроение, жанр). Это имеет ключевое значение в обработке аудиофайлов, т. к. информация, которую может получить в результате анализа CNN-сеть, является эффективным дополнением для поддержки определения и выстраивания новых данных RNN-сетью на основе полученных уже ею закономерностей.

Кроме того, отдельно стоит отметить, что работа CNN-сетей позволяет распознавать и отделять отдельные партии от общей аудиодорожки. Например, это применимо для изъятия вокальных партий из старых записей, создания а-капелл, но более всего это применимо для звукорежиссёров в области звукозаписи и звукосведения, что способно значительно упростить данные операции как в любительском, так и в профессиональном подходе.

Таким образом предоставляется возможность менять в композиции ритм, тональность или даже жанр целиком, добавляя или убирая целые партии музыкальных инструментов или эффектов. Нейросеть способна определять какие звуковые эффекты, будь то реверберация (reverb) или дилэй (delay), нужно

Дневник науки | [www.dnevnika.ru](http://www.dnevnika.ru) | СМИ ЭЛ № ФС 77-68405 ISSN 2541-8327

применить к композиции, какие музыкальные паттерны в контексте произведения добавить.

Применяя разные типы нейросетей вместе, музыканты или рядовые пользователи ИИ-сервисов получают возможность обработки музыкальных фраз без необходимости обращения к музыкальной теории, однако, это не означает полное отсутствие необходимости в её понимании. Не стоит забывать, что нейросеть следует использовать как инструмент, а не замену человеческому мыслительному процессу.

Тем не менее уже сейчас стоит отметить, что ИИ понимает различия между музыкальными жанрами, понимает их основы, благодаря которым мы сами определяем к какому жанру принадлежит та или иная композиция. [2,3]

Для решения второй задачи рассмотрим наиболее популярные сервис-редакторы.

**Musicfy** – платформа для создания музыки с помощью искусственного интеллекта. Инструмент предназначен для музыкантов, продюсеров и любителей музыки, подходит для разных уровней навыков. Некоторые функции:

- доступ к вокалу без авторских прав, чтобы обогатить треки новыми звуками;
- создание собственного ИИ. Можно загрузить вокал, чтобы создать модель искусственного интеллекта, точно воспроизводящую голос пользователя;
- разделители стеблей. Позволяют изолировать партии внутри песни, например ударные, вокал и бас, для более точного контроля;
- текст AI в музыку. Написанные слова и эмоции преобразуются в уникальные песни;
- голоса с использованием искусственного интеллекта. Ремейки песен с разными голосами.

**Jukebox** – нейросеть для создания музыки, выпущенная компанией OpenAI в 2020 году.

Некоторые особенности Jukebox:

- работа с настоящими аудиофайлами. В отличие от традиционных систем, которые генерируют музыку, опираясь на ноты или символы, Jukebox создаёт законченные музыкальные произведения с характерными для живых исполнений эмоциями и акцентами;
- создание музыки в различных жанрах. Jukebox может генерировать песни, например, в стиле рок, хип-хоп и джаз;
- имитация мелодии, ритма и звучания разных инструментов. Также нейросеть способна создавать вокал, который будет звучать вместе с музыкой.

Для работы Jukebox необходимы навыки программирования, а также мощное оборудование и сервер. Кроме того, процесс генерации может занимать много времени: одна минута трека генерируется несколько часов.

**SoftVC VITS Singing Voice Conversion (So-VITS-SVC)** – нейросеть на базе искусственного интеллекта, которая способна имитировать голоса певцов и создавать новые песни с их голосами.

Для обучения пользователи извлекают дорожку вокала из записей известных музыкантов, загружают сэмплы в программу, а затем генерируют вокальные партии заданного исполнителя.

Некоторые возможности So-VITS-SVC:

- выбор голоса-источника из списка заранее подготовленных голосов или загрузка собственного файла;
- настройка высоты звучания генерируемого голоса;

- настройка динамики генерируемого голоса;
- настройка длительности звуков и смены темпа в ходе разговора/пения для будущего голоса;
- сохранение результатов генерации в WAV-файл.

На основе имеющейся информации о представленных нейросетях была создана таблица ниже для сравнения их базовых критериев для начинающего пользователя (табл.1)

Таблица 1 – Сравнение базовых критериев пользования основных музыкальных нейросетей начинающим пользователем

	Musicfy	Jukebox	So-VITS-SVC
модель распространения	коммерческая	некоммерческая	некоммерческая
ресурсоёмкость	низкая	крайне высокая	средняя
настраиваемость	детальный контроль отсутствует	высокая	высокая
интерфейс	удобный web-интерфейс	UI отсутствует, взаимодействие через код и скрипты	смешанный
документация и сообщество	имеется	базовая, требуются технические знания	широкое сообщество

Таким образом следует, что Musicfy удобен и прост для быстрого использования с минимальными требованиями, но с ограниченной гибкостью. Jukebox — мощный инструмент для продвинутых пользователей с высоким уровнем ресурсов и навыков. So-VITS-SVC — оптимальный вариант для голосового синтеза с хорошей настраиваемостью и открытостью при умеренных требованиях.

Вполне вероятно, что нейросети в контексте музыки ждёт та же судьба, что и некогда электронные инструменты, музыкальное программное обеспечение: по началу оно будет использоваться в основном энтузиастами, формируя сообщество вокруг себя, но вскоре начнёт перениматься крупными студиями, профессиональным бизнесом в сфере музыки и аудио, что превратит нейросети

в новый, эффективный и признанный инструмент при работе в звукозаписи, звукорежиссуре и других специальностях области.

### **Библиографический список:**

1. Агазода Р. Нейросети и песни: как заставить Канье Уэста петь «Гражданскую Оборону» [Электронный ресурс]. — Режим доступа — URL: <https://tproger.ru/articles/nejroseti-pesni> (Дата обращения 28.01.2026)
2. Антонян К. Г. КУЛЬТУРА МЭШАП: СТРАТЕГИИ АПРОПРИАЦИИ / К. Г. Антонян, Н. А. Соколова, О. Н. Левшина // Вестник Московского государственного лингвистического университета. Гуманитарные науки: сборник научных работ по искусствоведению 2022 года – Москва: Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный лингвистический университет», 2022. – С. 144–150. – ISSN 2542–2197.
3. Дмитриев, А. В. Использование искусственного интеллекта для генерации музыкальных произведений / А. В. Дмитриев, Е. Д. Пойманова // Радиотехнические, оптические и биотехнические системы. Устройства и методы обработки информации: Сборник докладов. Четвертая Всероссийская научная конференция, Санкт-Петербург, 10–17 апреля 2023 года. – Санкт-Петербург: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2023. – С. 185–187. – EDN FIXBXK.
4. Котовенко, А. Ю. Применение искусственного интеллекта в музыкальной индустрии / А. Ю. Котовенко, Е. А. Минина // Цифровизация экономики: направления, методы, инструменты : Сборник материалов VII всероссийской научно-практической конференции, Краснодар, 15–19 января 2025 года. – Краснодар: ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина», 2025. – С. 151–154. – EDN QRFICT.