

Сравнительный анализ бесплатных версий нейросетей Music GPT и SUNO для создания песен

Пайдершев Алексей Константинович

*Приамурский государственный университет имени Шолом-Алейхема
Студент*

Аннотация

Цель данного исследования заключается в демонстрации возможностей бесплатных версий нейросетевых платформ SUNO и Music GPT, предназначенных для поэтапной генерации музыкальных произведений на основе готовых текстов песен и текстовых описаний. Выбор данных нейросетевых платформ обусловлен простотой использования и наличием бесплатного пробного периода. Такой подход позволяет провести сравнительный анализ двух инструментов, на основании которого можно сделать выбор в сторону платформы, которая больше подходит пользователю.

Ключевые слова: песня, нейросеть, SUNO, Music GPT.

A comparative analysis of the free versions of Music GPT and SUNO neural networks for creating songs

Paydershev Alexey Konstantinovich

*Sholom-Aleichem Priamursky State University
Student*

Abstract

The purpose of this study is to demonstrate the capabilities of the free versions of the SUNO and Music GPT neural network platforms, which are designed for the step-by-step generation of musical compositions based on pre-existing song lyrics and text descriptions. The choice of these neural network platforms is based on their ease of use and the availability of a free trial period. This approach allows for a comparative analysis of the two tools, which can be used to determine which platform is more suitable for the user.

Keywords: song, neural network, SUNO, Music GPT.

1 Введение

1.1 Актуальность

Актуальность данного исследования обусловлена стремительным развитием нейросетевых технологий, которые кардинально трансформируют подходы к созданию аудиального контента. В эпоху стремительного развития искусственного интеллекта генерация музыки на основе текстовых промптов приобретает стратегическое значение для демократизации творчества, позволяя авторам без профессионального музыкального образования

создавать полноценные треки. Бесплатные версии платформ SUNO и Music GPT выделяются простотой использования и пробным доступом, однако отсутствие сравнительных исследований их возможностей препятствует информированному выбору пользователей.

1.2 Обзор исследования

К. Ю. Пауэр, А. А. Кузнецов использовали нейросети как инструмент генерации художественных образов на примере клипа «Убегает весь мир» на песню Егора Летова «Без меня» [1]. Способы взаимодействия композитора с искусственным интеллектом и проблема авторства сгенерированного нейросетью произведения продемонстрировал В. Е. Григорьев [2]. А. Р. Давлетгареева и К. А. Едкова проанализировали нейронную сеть для генерации изображений на основе текста песен с применением моделей OpenAI и CLIP [3]. Генерацию музыки с использованием двунаправленной рекуррентной нейронной сети осуществили Е. Д. Мосин, Ю. С. Белов [4]. И. А. Батраева, А. Д. Нарцев, А. С. Лезгян классифицировали тексты с использованием сверточной нейронной сети на основе векторного представления слов [5].

1.3 Цель исследования

Целью исследования является демонстрация возможностей бесплатных версий нейросетевых платформ SUNO и Music GPT для поэтапной генерации музыкальных произведений на основе готовых текстов песен и текстовых описаний, провести сравнительный анализ этих инструментов с учетом их простоты использования и наличия бесплатного пробного периода, а также обосновать выбор наиболее подходящей платформы для конкретного пользователя.

2 Материалы и методы

В качестве основных инструментов исследования применялись нейросетевые платформы SUNO и Music GPT, реализующие технологию генерации музыкальных произведений через текстовые описания. Такой метод обеспечивает высокую точность в управлении композиционными, стилистическими и аудиальными параметрами результирующего музыкального контента. Для создания текста песни использована мультимодальная нейросетевая платформа Perplexity.

3 Результаты

Первым шагом исследования стало написание текста песни с помощью мультимодальной нейросетевой платформы Perplexity. Промт позволяет создать текст песни в духе русскоязычного пост-панка и лирики, близкой к настроению советской и российской рок-группы «КИНО», в котором прослеживаются простые образы, которые часто использовал в своих песнях Виктор Цой – борьба, путь, решимость.

Промт: Придумай текст песни про осень в духе русскоязычного пост-панка и лирики, близкой к настроению советской и российской рок-группы «КИНО». Используй простые образы – борьба, путь, решимость. (Come up with a song about autumn in the spirit of Russian-language post-punk and lyrics that are similar to the mood of the Soviet and Russian rock band KINO. Use simple imagery such as struggle, path, and determination.)

Текст песни:

Куплет 1

Осень стучится в закрытые двери,
Дворник рисует круги на ветру.
Мы так устали играть в ваши веры,
Но по привычке опять не умрём к утру.

Солнце за городом медленно тает,
Сотни витрин отражают глаза.
Каждый из нас здесь чего-то теряет,
Но остаётся идти, даже если нельзя.

Припев

А за окном опять идёт дождь,
Словно стирает наш вчерашний след.
Если ты всё равно уйдёшь,
Значит, у нас с тобой больше нет побед.

Осень сожжёт ненужный хлам,
Ветер раздаст по карманам ноль.
Если весь мир идёт по вам,
Вы выбирайте движение, а не боль.

Куплет 2

Город гудит, как пустая станция,
Поезд прошёл, но никто не зашёл.
Время меняет направление концов,
Кто не успел — тот и не нашёл.

Мы поднимаемся в старый подъезд,
Где за таблицами спрятан свет.
Каждый решает, что правда есть,
Если за правду приходит рассвет.

Припев

А за окном опять идёт дождь,
Смазывая чёрный контур домов.
Если ты всё равно уйдёшь,
Не оставляй за собой лишних слов.

Осень считает чужие долги,
Сбрасывает листья, как старый страх.
Если мы так и не стали большими,
Значит, ещё не закончился наш крах.

Бридж

И кто-то снова включает свет,
Пряча глаза в сигаретный дым.
Там, где вопросов давно уже нет,
Каждый становится сам себе чужим.

Финальный припев

А за окном опять идёт дождь,
Словно проверяет на прочность стекло.
Если ты всё равно уйдёшь,
Значит, всё это было не зря, но прошло.

Осень рисует на небе знак,
Серой чертой отделяя «вчера».
Если весь мир идёт на нас,
Мы всё равно выходим из двора.

Вторым шагом стало написание промта для нейросетевых аудиальных платформ.

Промт: Создайте песню в стиле легендарной советской рок-группы «Кино» и голосом Виктора Цоя. Используйте постпанк с элементами новой волны, минималистичные гитарные риффы, ритмичную драм-машину вместо живых барабанов, простой бас и атмосферные синтезаторы для создания ощущения 80-х в СССР. Голос: мужской баритональный, как у Виктора Цоя — с металлическим призвуком, мелким вибрато, грудным резонатором на низких нотах, минимальной артикуляцией, округляющимися словами и мощной энергетикой без лишних вокальных приёмов, подходящей для стадионного звучания. Текст: на русском языке, в поэтическом стиле Цоя — короткие пронзительные строки о жизни, ночи, звёздах, переменах, одиночестве, протесте или повседневной реальности (вдохновляясь «Группа крови», «Звезда по имени Солнце», «Пачка сигарет» или «Хочу перемен»). Структура: 2-3 куплета, припев, бридж; ритм 120-140 BPM, меланхоличная, но драйвовая атмосфера с нарастающим напряжением. Длина: 2-3 минуты, энергичный припев для хука. Сделайте так, чтобы звучало аутентично, как будто Цой поёт новую песню «Кино» в 2026 году. (Create a song in the style of the legendary Soviet rock band Kino and the voice of Viktor Tsoi. Use post-punk with elements of new wave, minimalist guitar riffs, a rhythmic drum machine instead of live drums, a simple bass, and atmospheric synthesizers to create the feeling of the 1980s in the USSR. Voice: male baritone, like Viktor Tsoi's, with a metallic undertone, subtle vibrato, a chesty resonator on low notes, minimal articulation,

rounded words, and a powerful energy without unnecessary vocal techniques, suitable for stadium sound. Lyrics: in Russian, in Tsoi's poetic style, with short, poignant lines about life, the night, the stars, change, loneliness, protest, or everyday reality (inspired by «Blood Type», «Star Named Sun», «Pack of Cigarettes», or «I Want Change»). Structure: 2-3 verses, chorus, bridge; 120-140 BPM rhythm, melancholic but driving atmosphere with increasing tension. Length: 2-3 minutes, energetic chorus for the hook. Make it so that)

Следующий шаг – загрузка промта в аудиальные нейросетевые платформы SUNO и Music GPT. На данном этапе Music GPT проявил себя лучше, так как количество знаков в промте, который платформа позволяет загрузить, не ограничено, в SUNO же такой возможности нет. Поэтому промт для SUNO пришлось сократить.

Промт для SUNO: Создайте песню в стиле «Кино» голосом Виктора Цоя: постпанк 80-х с минималистичными гитарами, драм-машиной, простым басом и синтезаторами, 120-140 BPM. Голос: баритональный, металлический с вибрато, грудной, энергичный без излишеств, как у Цоя. Текст на русском: короткие строки о ночи, переменах, одиночестве (как в «Группа крови» или «Звезда по имени Солнце»); 2 куплета, припев, бридж; 2-3 минуты, меланхолично-драйвовая. (Create a Kino-style song with Victor Tsoi's voice: 80s post-punk with minimalistic guitars, drum machine, simple bass and synthesizers, 120-140 BPM. Voice: baritone, metallic with vibrato, chesty, energetic without excesses, like Tsoi. Lyrics in Russian: short lines about the night, change, and loneliness (like in «Blood Type» or «A Star Named Sun»); 2 verses, chorus, and bridge; 2-3 minutes, melancholic and driving.)

Ссылка на материалы, сгенерированные платформой Music GPT <https://disk.yandex.ru/d/6ysEsNh8z-Y-Pg>.

Ссылка на материалы, сгенерированные платформой SUNO https://disk.yandex.ru/d/gbJprTyA4_5b7Q.

После загрузки промта и текста песни в обе платформы, нейросети сгенерировали по два варианта песни с разными аранжировками и голосами.

Первое, что бросается в глаза – ни в одном варианте голос Виктора Цоя воссоздать не удалось. Хотя тембрально варианты, предложенные Music GPT намного ближе к оригинальному голосу, манера исполнения сильно отличается.

Голосовое сопровождение песен, которые сгенерировала SUNO и вовсе не похожи на оригинал, ни тембрально, ни манерой исполнения. С этой точки зрения Music GPT явно выигрывает.

Второй аспект сравнения – аранжировка. В первом варианте, который предоставила SUNO, вопреки промту, отсутствуют некоторые инструменты – гитары и бас-гитара. Присутствует только драм-машина и синтезатор. Во второй версии так же отсутствует бас-гитара и ритм-гитара, хотя в финале песни отчётливо слышно соло-гитару. Однако эта гитара не соответствует стилю «КИНО» - подобных партий соло-гитарист группы Юрий Каспарян никогда не использовал. Обе аранжировки современные, как и было указано в промте, но отличительный стиль группы не прослеживается.

Инструментал вариантов песен, которые создала Music GPT, намного ближе к оригинальным аранжировкам «КИНО» - все инструменты, которые были указаны в промте отчётливо слышны, а партии этих инструментов достаточно похожи на исполнение группы. Однако современными их назвать сложно.

Так же стоит отметить, что функционал бесплатных версий обеих платформ ограничен, но идентичен – все бесплатные инструменты одной платформы совпадают со второй. При этом в SUNO каждый день можно генерировать до 10 проектов по 2 трека, в Music GPT же весь бесплатный период ограничен 5 проектами по 2 трека, а если их редактировать – ещё меньше. В платных версиях этих платформ есть возможность добавления референсов – голосовых и инструментальных. При их помощи возможно воссоздать голос, манеру речи и прочие отличительные черты исполнителя.

4 Вывод

В результате процесса генерации аудиального контента на бесплатных версиях платформ Music GPT и SUNO, основанного на текстовых описаниях были получены четыре самостоятельных музыкальных произведения. Созданные произведения не совпадают с желаемым результатом, однако если отсутствует цель воссоздать похожий голос или аранжировку, треки, полученные на основе текстовых описаний, обладают музыкальной целостностью и могут считаться законченными музыкальными произведениями. Это открывает перспективы для создания оригинальных музыкальных проектов, но не подходит для воссоздания чужих голосов или аккомпанемента.

Библиографический список

1. Пауэр К. Ю., Кузнецов А. А. Нейросети как инструмент генерации художественных образов на примере клипа «Убегает весь мир» на песню Егора Летова «Без меня» //Русская рок-поэзия: текст и контекст. 2021. №. 21. С. 164-176. URL: <https://cyberleninka.ru/article/n/neyroseti-kak-instrument-generatsii-hudozhestvennyh-obrazov-na-primere-klipa-ubegaet-ves-mir-na-pesnyu-egora-letova-bez-menya>
2. Григорьев В. Е. Способы взаимодействия композитора с искусственным интеллектом и проблема авторства сгенерированного нейросетью произведения //Вестник Академии русского балета им. АЯ Вагановой. 2024. №. 2 (91). С. 137-153. URL: <https://cyberleninka.ru/article/n/sposoby-vzaimodeystviya-kompozitora-s-iskusstvennym-intellektom-i-problema-avtorstva-sgenerirovannogo-neyrosetyu-proizvedeniya>
3. Давлетгареева А. Р., Едкова К. А. Нейронная сеть для генерации изображений на основе текста песен с применением моделей OpenAI и CLIP //Электронные библиотеки. 2023. Т. 26. №. 4. С. 437-455. URL: <https://rdl-journal.ru/article/view/790>
4. Мосин Е. Д., Белов Ю. С. Генерация музыки с использованием

- двунаправленной рекуррентной нейронной сети //Научное обозрение. Технические науки. 2023. Т. 1. С. 10-14. URL: https://s.science-engineering.ru/pdf/2023/2023_1.pdf#page=10
5. Батраева И. А., Нарцев А. Д., Лезгян А. С. Классификация текстов с использованием сверточной нейронной сети на основе векторного представления слов //Информационные технологии и математическое моделирование (ИТММ-2019). 2019. С. 114-119. URL: <https://vital.lib.tsu.ru/vital/access/services/Download/vtls:000661768/SOURCE1?view=true#page=114>