



Учитель Начальной

**Курс нейросетей для
композитора музыки**

«Автор»

2026

Начальной У. Ш.

Курс нейросетей для композитора музыки / У. Ш. Начальной —
«Автор», 2026

Практический курс для композиторов, продюсеров и саунд-дизайнеров, которые хотят освоить нейросети как полноценный творческий инструмент. Книга последовательно проводит читателя от базовых принципов машинного обучения до профессиональной работы с современными генеративными системами. Разбираются архитектуры RNN, LSTM, Transformer, GAN, VAE и диффузионные модели. Подробно рассматриваются инструменты Magenta, Suno, Udio, Stable Audio и AIVA. Читатель учится готовить MIDI-данные, писать сильные промпты, управлять структурой и эмоциями генерации, дообучать модели на своём материале и встраивать ИИ в рабочий процесс DAW. Отдельные главы посвящены оценке качества, авторским правам и этике. Курс завершается практическим проектом и дорожной картой развития на 3–6 месяцев. Издание сочетает теорию с большим количеством примеров, шпаргалок и рекомендаций, помогая сохранить авторский голос при работе с искусственным интеллектом.

© Начальной У. Ш., 2026

© Автор, 2026

Содержание

Глава 1. Введение: нейросети как новый инструмент композитора	5
1.1. Зачем композитору нейросети	5
1.2. Краткая история ИИ в музыке	6
1.3. Что вы получите из этого курса	7
1.4. Как работать с курсом	8
Глава 2. Основы машинного обучения для музыкантов	9
2.1. Что такое нейронная сеть	9
2.2. Основные понятия	10
2.3. Почему музыка - сложная задача для ИИ	11
2.4. Полезные аналогии для музыканта	12
Глава 3. Как компьютер слышит и записывает музыку	13
3.1. Два мира: символьная музыка и аудио	13
3.2. MIDI и piano-roll	14
3.3. Спектрограммы и mel-spectrogram	15
3.4. Токенизация музыки	16
Глава 4. Рекуррентные сети и LSTM: первые генеративные модели	17
4.1. Идея RNN	17
Конец ознакомительного фрагмента.	18

Учитель Начальной

Курс нейросетей для композитора музыки

Глава 1. Введение: нейросети как новый инструмент композитора

1.1. Зачем композитору нейросети

Нейросети - это не замена композитору, а мощный инструмент, расширяющий творческие возможности. Они помогают:

- преодолевать творческий блок (когда не идёт мелодия или гармония);
- быстро прототипировать идеи (за минуты получить десятки вариантов);
- исследовать новые гармонические и ритмические пространства;
- генерировать фоновую музыку, аранжировки, вариации и сырой материал;
- создавать музыку под конкретные задачи (игры, реклама, подкасты) с минимальными затратами времени.

Исторически композиторы всегда использовали технологии: от фортепиано и органа до синтезаторов, сэмплеров и DAW. Нейросети - логичное продолжение этой линии. Разница лишь в том, что теперь машина может предлагать не только звуки и тембры, но и целые музыкальные идеи, структуры и даже готовые треки.

Пример из практики. Композитор пишет саундтрек к короткометражке. Вместо того чтобы часами подбирать атмосферу, он генерирует в Stable Audio или AIVA 8-10 вариантов ambient-трека по описанию сцены, выбирает два лучших, дорабатывает их в DAW и получает готовый материал за вечер вместо двух дней.

1.2. Краткая история ИИ в музыке

Алгоритмическая композиция существует давно:

-

XVIII век - Музыкальные кости Моцарта (Musikalisches Würfelspiel) - генерация менуэтов случайным образом.

-

1950-60-е - первые компьютерные эксперименты (Lejaren Hiller, Illiac Suite).

-

1980-90-е - rule-based системы и марковские цепи.

-

2000-е - Hidden Markov Models и первые нейронные подходы.

Настоящий прорыв произошёл в 2010-х с глубоким обучением:

-

2015-2016 - первые успешные LSTM-модели для мелодий.

-

2016 - запуск Google Magenta.

-

2019 - OpenAI MuseNet (Transformer), генерация до 4 минут музыки с 10 инструментами.

-

2020-2022 - Music Transformer, MusicVAE, Jukebox, первые диффузионные эксперименты.

-

2022-2023 - MusicLM, Riffusion, текстово-музыкальные модели.

-

2023-2026 - Suno, Udio, Stable Audio: генерация полноценных песен с вокалом студийного качества.

1.3. Что вы получите из этого курса

После прохождения курса вы будете:

- понимать принципы работы основных генеративных архитектур (RNN/LSTM, Transformer, GAN, VAE, диффузия);
- уверенно пользоваться готовыми сервисами (Suno, Udio, Stable Audio, AIVA, Magenta);
- готовить MIDI- и аудио-данные и дообучать модели под свой стиль;
- интегрировать ИИ в свой композиторский workflow без потери авторского голоса;
- ориентироваться в вопросах авторских прав и этики;
- иметь готовую дорожную карту дальнейшего развития.

1.4. Как работать с курсом

Рекомендуемый подход:

Читайте главу и сразу пробуйте описанные инструменты.

Ведите дневник экспериментов: что сгенерировали, какие промпты сработали, что пришлось править вручную.

Не стремитесь сразу к идеальной автономной генерации. Цель - получить материал, который интересно дорабатывать.

После глав 7-8 и 11-13 сделайте небольшой реальный проект (трек, саундтрек, заставка).

Глава 2. Основы машинного обучения для музыкантов

2.1. Что такое нейронная сеть

Нейронная сеть - это математическая модель, вдохновлённая биологическими нейронами. Она состоит из слоёв нейронов, которые преобразуют входные данные (например, последовательность нот) в выходные (следующие ноты или целую композицию).

Обучение происходит путём минимизации ошибки предсказания на большом количестве примеров. Модель учится статистическим закономерностям музыки: какие ноты часто следуют друг за другом, как строятся фразы, какие гармонии типичны для стиля, как развивается динамика.

Аналогия для музыканта. Представьте, что вы слушали тысячи произведений одного стиля и у вас сформировалась интуиция что обычно бывает дальше. Нейросеть делает то же самое, но на математическом уровне и в гораздо больших масштабах.

2.2. Основные понятия

-

Датасет -

коллекция примеров (MIDI-файлы, аудиозаписи), на которых учится модель. Качество и разнообразие датасета критически важны.

-

Признаки (features) -

числовое представление музыки (высота тона, длительность, velocity, chroma, MFCC и т.д.).

-

Обучение с учителем / без учителя -

в генеративной музыке чаще используется unsupervised learning (предсказание следующего токена).

-

Overfitting -

когда модель слишком хорошо запоминает обучающие данные и плохо обобщает (начинает копировать куски треков).

-

Underfitting -

модель слишком простая и не улавливает закономерности.

-

Гиперпараметры -

настройки архитектуры и обучения (размер слоёв, learning rate, batch size, температура сэмплирования).

-

Loss (функция потерь) -

мера того, насколько модель ошибается. При обучении мы её минимизируем.

-

Epoch -

один полный проход по всему датасету.

2.3. Почему музыка - сложная задача для ИИ

Музыка обладает:

- иерархической структурой (ноты - мотивы - фразы - разделы - форма);
- долгосрочными зависимостями (реприза через 2 минуты);
- полифонией (несколько голосов одновременно);
- субъективной оценкой качества (что красиво - вопрос вкуса и культуры);
- многомерностью (высота, ритм, тембр, динамика, артикуляция).

Модель должна одновременно учитывать локальную гармонию и глобальную драматурию. Это намного сложнее, чем, например, классификация изображений.

2.4. Полезные аналогии для музыканта

-

Температура сэмплирования

примерно равна тому, насколько свободно импровизирует музыкант. Низкая температура (0.6-0.8) - более предсказуемо и правильно. Высокая (1.1-1.4) - более неожиданно, но рискованно.

-

Context length

- насколько далеко модель помнит предыдущие ноты.

-

Fine-tuning

- когда вы берёте готового музыканта и учите его играть именно в вашем стиле.

Глава 3. Как компьютер слышит и записывает музыку

3.1. Два мира: символьная музыка и аудио

Существует два принципиально разных подхода к представлению музыки.

1. Символьное (MIDI, piano-roll, ABC-нотация, MusicXML)

- дискретное представление нот, ритма, инструментов;
- удобно для композиции, редактирования, точного контроля;
- модели: Magenta RNN/VAE, Music Transformer, MuseNet.

2. Аудио (waveform, спектрограммы)

- непрерывный сигнал;
- позволяет генерировать тембр и вокал;
- сложнее контролировать структуру и отдельные ноты;
- модели: WaveNet, MusicLM, Stable Audio, Suno, Udio.

Практический вывод. Если вам нужен контроль над нотами и возможность править в нотном редакторе или DAW - работайте с MIDI-моделями. Если нужен готовый звук из коробки (особенно с вокалом) - аудио-модели.

3.2. MIDI и piano-roll

MIDI хранит события: Note On/Off, pitch (0-127), velocity, control changes, tempo, time signature.

Piano-roll - это матрица время на высоту тона, где ячейка содержит информацию о ноте (часто binary или velocity). Именно piano-roll чаще всего используют как вход и выход для символьных моделей.

Пример представления простой мелодии: Доля 1: C4 (pitch 60), Доля 2: E4 (pitch 64), Доля 3: G4 (pitch 67), Доля 4: C5 (pitch 72). В токенизированном виде это может выглядеть как последовательность: NOTE_ON_60, TIME_SHIFT_12, NOTE_OFF_60, NOTE_ON_64 и так далее.

3.3. Спектрограммы и mel-spectrogram

Для работы с аудио сигнал преобразуют в спектрограмму с помощью коротковременного преобразования Фурье (STFT). Mel-спектрограмма учитывает особенности человеческого слуха (мы лучше различаем низкие частоты) и является стандартным представлением для современных аудио-моделей.

Многие модели генерируют именно mel-спектрограмму, а затем используют нейронный вокодер (HiFi-GAN, BigVGAN и др.), чтобы превратить её в waveform.

3.4. Токенизация музыки

Современные Transformer-модели превращают музыку в последовательность дискретных токенов - по аналогии с токенами в языковых моделях (GPT). Токены могут кодировать ноту (pitch + octave), инструмент, длительность / time-shift, velocity / динамику, специальные события (начало такта, смена тональности, pedal). Чем лучше продумана схема токенизации, тем лучше модель улавливает музыкальную структуру.

Глава 4. Рекуррентные сети и LSTM: первые генеративные модели

4.1. Идея RNN

Рекуррентные нейронные сети (RNN) обрабатывают последовательности, сохраняя память о предыдущих шагах в скрытом состоянии. Это идеально подходит для музыки, которая разворачивается во времени. На каждом шаге сеть получает текущую ноту (или событие) и предыдущее скрытое состояние, выдаёт предсказание следующей ноты и обновлённое состояние.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.