

ПРИМЕНЕНИЕ ГЕНЕРАТИВНЫХ НЕЙРОННЫХ СЕТЕЙ В СФЕРЕ ДИЗАЙНА ОДЕЖДЫ

APPLICATION OF GENERATIVE NEURAL NETWORKS IN FASHION DESIGN

А. О. ЯБЛОНСКАЯ
A. O. YABLONSKAYA

Белорусский государственный университет
Минск, Беларусь
Belarussian state university
Minsk, Belarus
e-mail: anna.yablonskaya2002@gmail.com

В данном докладе рассматривается применение генеративных нейронных сетей для создания изображений одежды по запросам пользователя. Представлена общая информация о генеративных моделях, а также информация об используемой для построения модели, архитектуре Kandinsky. В рамках доклада также представлены результаты обучения данной модели.

Ключевые слова: система онлайн-стилиста; генеративная модель; генерация изображений; модель Kandinsky; архитектура модели; обучение модели.

This report discusses the use of generative neural networks to generate clothing images based on user requests. In the report you can find general information about generative models as well as the information on the Kandinsky architecture, used to build a model. Examples of the training results are provided as well.

Keywords: online-stylist system; generative model, image generation; Kandinsky model; model architecture; model training.

ВВЕДЕНИЕ

В современном мире мода и стиль играют важную роль в жизни людей, определяя не только внешний вид, но и самоидентификацию. В последние годы индустрия моды стала свидетелем революционного сдвига благодаря достижениям в области искусственного интеллекта и машинного обучения. Стремительное развитие данной сферы привело к созданию нейронных сетей, которые, обучаясь на огромных

объёмах данных, позволяют дизайнерам создавать инновационную, индивидуальную и устойчивую моду, как никогда раньше. Подражая человеческому творчеству и распознаванию образов, нейронные сети могут создавать новые дизайны, оптимизировать текстильные узоры и даже предсказывать будущие тенденции, преобразуя способ воображения, создания и восприятия дизайна в сфере моды.

Стали популярны системы онлайн-стилистов, использующие методы машинного обучения. Они предлагают инновационные решения для персонализированного подхода в выборе одежды и аксессуаров. Эти технологии способны анализировать предпочтения пользователей, типов их телосложения, комбинацию цветов и форм, а также учитывать актуальные тренды, что значительно упрощает процесс поиска подходящих предметов одежды и создание стильных образов. Эти платформы на базе искусственного интеллекта позволяют людям получать индивидуальные предложения по одежде без необходимости в личном стилисте, делая моду более доступной. К тому же, поскольку в последние годы значительную популярность и развитие приобрели генеративные нейронные сети, их начали активно внедрять в системы онлайн-стилистов.

Использование генеративных нейронных сетей (*GNN*) в дизайне одежды стремительно набирает популярность. В системе онлайн-стилиста генеративные нейронные сети (*GNN*) могут использоваться для улучшения возможностей персонализации и дизайна путём создания индивидуальных предложений по одежде и даже создания новых комбинаций нарядов. Также одной из подзадач в создании онлайн-стилиста является генерация новых изображений на основе описания предмета и исходной картинке предмета. Такая генерация даёт возможность клиентам видеть их идеи и пожелания в визуальной форме. Эта возможность может позволить системе стилистов предлагать уникальные дизайны по запросу пользователя, основываясь на их предпочтениях и запросах. Что также улучшает процесс покупок, поскольку далее такие изображения можно будет применять для поиска схожих товаров в доступных вариантах. Более того, пользователи могут активно участвовать в создании и настройке своих образов, что делает взаимодействие с онлайн-стилистом более увлекательным и персонализированным.

В данной статье будут представлены примеры применения генеративных нейронных сетей в системах онлайн стилистов, а также

представлены результаты построенной в рамках статьи генеративной модели для создания изображений одежды.

ГЕНЕРАТИВНЫЕ МОДЕЛИ

Генеративные модели (*Generative Models*) – это класс искусственных нейронных сетей, разработанных для генерации новых данных, которые напоминают заданный набор данных. Эти модели обычно используются для таких задач, как синтез изображений, генерация текста, сочинение музыки и других творческих приложений. Они отличаются от традиционных нейронных сетей, которые обычно фокусируются на таких задачах, как классификация или регрессия. Вместо этого генеративные модели изучают базовое распределение данных и используют эти знания для создания новых, похожих экземпляров.

Одним из самых популярных типов генеративных моделей является генеративно-сопоставительная сеть (*Generative Adversarial Network, GAN*), представленная Яном Гудфеллоу в 2014 году. GAN состоят из двух нейронных сетей: генератора и дискриминатора. Роль генератора заключается в создании поддельных данных (например, изображения), в то время как роль дискриминатора заключается в различении реальных и поддельных данных. Эти две сети обучаются вместе в своего рода «игре», где генератор пытается улучшить свои выходные данные, чтобы обмануть дискриминатор, а дискриминатор работает над тем, чтобы стать лучше в обнаружении подделок. Этот состязательный процесс приводит к тому, что генератор со временем выдаёт более реалистичные данные.

Другим важным подходом являются вариационные автокодировщики (*Variational Autoencoder, VAE*), которые основаны на другой концепции. VAE нацелены на моделирование распределения данных, обучаясь кодировать данные в скрытое пространство (сжатое, абстрактное представление данных), а затем декодировать их обратно в исходное пространство данных. Представление скрытого пространства позволяет генерировать новые данные путём выборки точек из этого пространства и их декодирования [2].

ПРИМЕРЫ СИСТЕМ ОНЛАЙН-СТИЛИСТОВ, ПРИМЕНЯЮЩИХ ГЕНЕРАТИВНЫЕ МОДЕЛИ

The Yes – это торговая платформа на основе искусственного интеллекта, которая использует машинное обучение и генеративные сети для предоставления персонализированных рекомендаций по одежде

и аксессуарам на основе предпочтений пользователя. Она учится на поведении пользователя и интегрирует генеративный искусственный интеллект, чтобы предлагать стили, соответствующие индивидуальным вкусам. На рис. 1 продемонстрированы образы, сгенерированные на данной платформе (Прим.14, рис.1).

Zalando – крупный европейский интернет-магазин модной одежды, использует различные технологии ИИ, включая генеративные сети, для создания персонализированного опыта покупок модной одежды. Их «Fashion Companion» использует ИИ, чтобы помогать пользователям находить новые товары на основе их индивидуальных предпочтений и прошлых покупок. Также применяется технология OpenAI, которая позволяет клиентам задавать вопросы, используя собственные модные термины и слова, помогая им ориентироваться в широком ассортименте Zalando более интуитивно понятным способом. На рис. 2 показан пример взаимодействия пользователя с приложением Zalando (Прим.14, рис.2).

Fashion++ – это исследовательский проект, который использует GAN для создания новых предметов одежды и их комбинаций. Хотя это и не сервис, ориентированный на потребителя, он демонстрирует, как генеративные модели могут использоваться для создания новых предметов моды на основе конкретных подсказок по стилю или вводных данных от пользователей. Эту технологию можно применять в системах онлайн-стилистов, позволяя им создавать новые наряды или уникальные дизайны, соответствующие предпочтениям пользователей. На рис. 3 представлено изменение образов в зависимости от текущих трендов (Прим.14, рис.3).

Chic AI использует генеративные модели и глубокое обучение, чтобы помочь модным брендам и ритейлерам создавать персонализированный опыт покупок. ИИ генерирует рекомендации по одежде на основе поведения пользователей при просмотре, истории покупок и даже их модных предпочтений, обеспечивая высоко персонализированный опыт покупок.

Outfit Generator by AI Fashion. AI Fashion использует генеративные сети для создания виртуальных нарядов на основе вещей в шкафу пользователя. Система предлагает комбинации, которые могут подойти к его существующему гардеробу, предлагая креативные предложения по стилю на основе последних тенденций и личного вкуса пользователя (Прим.14, рис.4).

МОДЕЛЬ KANDINSKY

Как можно заметить зачастую генеративные модели используются платформами онлайн-стилистов для персонализированных покупок и создании образов на основе уже существующих продуктов. Поэтому для реализации дизайнов одежды, задуманных пользователем, необходимо создать генеративную модель, способную генерировать изображения на основе существующих изображений и текстовых запросов. Для создания такой модели можно применить готовую архитектуру модели Кандинский.

Модель Кандинский (или Kandinsky) является примером генеративной модели, разработанной для создания изображений на основе текстовых описаний. Она основана на архитектуре трансформеров и предназначена для генерации высококачественного визуального контента. Существует различные версии данной модели.

Для создания модели генерации новых изображений была выбрана версия Kandinsky 2.1 [1]. Кандинский 2.1 унаследовал лучшие практики от *Dall-E 2* и скрытой диффузии, но привнёс некоторые новые идеи. В качестве кодировщика текста и изображений он использует модель *CLIP* и априорное диффузионное изображение (сопоставление между скрытыми пространствами модальностей *CLIP*). Этот подход повышает визуальную производительность модели и открывает новые горизонты в смешивании изображений и манипулировании изображениями с помощью текста. Для диффузного отображения скрытых пространств мы используем преобразователь с `num_layers=20`, `num_heads=32` и `Hidden_size=2048`.

Модель была дообучена на датасете *FashionGen*, содержащем 32 528 изображений с описаниями. После обучения, модель можно применять как для генерации изображений отдельно по текстовому запросу, так и на основе каких-либо изображений и текстового запроса.

Примеры.

1. На рис. 5а приведено изображение, полученное из исходного изображения 5б (Прим.14, рис.5) по запросу: «Пиджак того же стиля».

2. На рис. 6 приведено изображение, сгенерированное по запросу «Прямые леопардовые джинсы» (Прим.14, рис.6).

Выводы

Использование технологий машинного обучения для создания онлайн-стилистов представляет собой революционный шаг в сфере

моды и стиля. Это пересечение технологий и искусства не только оптимизирует процесс дизайна, но и открывает бесконечные возможности для настройки, эффективности и инклюзивности. Развитие генеративных нейронных сетей начало переопределять роль технологий в моде, предлагая свежие перспективы и решения проблем, с которыми сталкиваются как дизайнеры, так и потребители. Генерация изображений на основе описания и исходных картинок не только облегчает пользователям процесс подбора одежды, но и предоставляет уникальные возможности для самовыражения. В статье была представлена информация о генеративных моделях, примеры платформ онлайн-стилистов, а также результаты обучения генеративной модели Kandinsky, на основе которой была построена модель для генерации изображений одежды с пользовательским дизайном на основе как отдельного текстового запроса, так и исходного изображения и текстового описания.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. *Кузнецов А.* Kandinsky 2.1 или Когда +0.1 значит очень много. URL: <https://habr.com/ru/companies/sberbank/articles/725282/>. (дата обращения: 25.03.2025).
2. *Belcic I.* What is a generative model? URL: <https://www.ibm.com/think/topics/generative-model>. (дата обращения: 25.03.2025).