ЛЕКЦИЯ 11

ТЕХНОЛОГИИ ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Основные понятия • определения • схемы • таблицы • графики • примеры

Преподаватель: проф.Ахмедиярова А.Т.

Структура и основные разделы

01 Понятие генеративного искусственного интеллекта

02 Генеративные и дискриминативные модели

03 Что значит «генерировать» данные

04 Основные классы генеративных моделей

05 Авторегрессионная генерация

06 Автокодировщики

07 Вариационные автокодировщики VAE

08 Условные генеративные модели

09 Генеративно-состязательные сети GAN

10 Диффузионные модели

11 GAN, VAE и диффузионные модели: сравнение

12 Генерация текста

Понятие генеративного искусственного интеллекта

Генеративный искусственный интеллект (Generative AI) — направление искусственного интеллекта, в котором модели создают новый контент на основе закономерностей, выявленных в обучающих данных.

Ключевая идея: Генеративный искусственный интеллект (Generative AI) — направление искусственного интеллекта, в котором модели создают новый

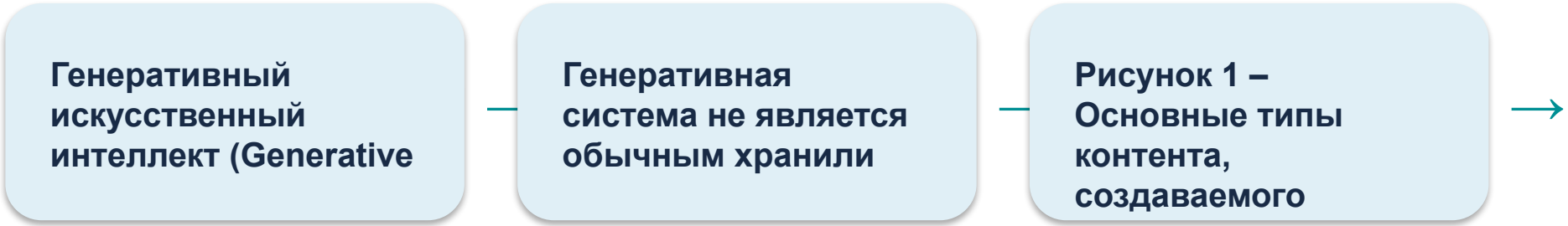
Понятие генеративного искусственного интеллекта — продолжение

В отличие от классификатора, который выбирает один из известных классов, генеративная модель формирует новый объект: текст, изображение, программный код, аудиосигнал, видеопоследовательность или другое представление данных. Генеративная система не является обычным хранилищем заранее подготовленных ответов. Результат формируется на основе параметров модели и текущего входного условия.

Понятие генеративного искусственного интеллекта — продолжение

Это условие может быть текстовым запросом, изображением, фрагментом аудио, кодом, структурированными данными или сочетанием нескольких типов информации. Рисунок 1 – Основные типы контента, создаваемого генеративным ИИ

Схема: Понятие генеративного искусственного интеллекта



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Генеративные и дискриминативные модели

Дискриминативные модели в первую очередь решают задачу различения объектов или прогнозирования целевого значения. Генеративные модели стремятся моделировать структуру данных так, чтобы создавать новые примеры, похожие по статистическим свойствам на обучающую выборку.

Сравнительная таблица: Генеративные и дискриминативные модели

Генеративные и дискриминативные модели	Дискриминативные модели в первую очередь решают задачу различения объектов или прогнозирования целевого значения. Г

Что значит «генерировать» данные

Генерация не означает случайное создание любого объекта. Модель формирует результат в соответствии с закономерностями обучающих данных и заданными условиями. Поэтому качество генерации зависит от состава обучающего набора, архитектуры модели, процедуры обучения и способа управления генерацией.

Ключевая идея: Генерация не означает случайное создание любого объекта. Модель формирует результат в соответствии с закономерностями обучающ

Что значит «генерировать» данные — продолжение

Например, модель генерации изображений получает текстовое описание и создаёт визуальный результат, соответствующий этому описанию. Языковая модель получает последовательность токенов и формирует продолжение, которое статистически согласуется с контекстом.

Авторегрессионная генерация

Авторегрессионная модель формирует последовательность поэлементно. На каждом шаге она оценивает вероятности возможных следующих элементов с учётом уже имеющегося контекста. Для текста элементом обычно является токен. Пусть модель получила фразу «Искусственный интеллект используется для».

Ключевая идея: Авторегрессионная модель формирует последовательность поэлементно. На каждом шаге она оценивает вероятности возможных следующ

Авторегрессионная генерация — продолжение

Она оценивает вероятности возможных продолжений и выбирает один из токенов, после чего повторяет процесс уже с расширенным контекстом. Выбор следующего элемента может быть детерминированным или вероятностным. Изменяя параметры генерации, можно получить более предсказуемые или более разнообразные результаты.

Автокодировщики

Автокодировщик состоит из двух основных частей: энкодера и декодера. Энкодер преобразует входные данные в компактное скрытое представление, а декодер восстанавливает исходные данные из этого представления. Обычный автокодировщик прежде всего решает задачу реконструкции.

Ключевая идея: Автокодировщик состоит из двух основных частей: энкодера и декодера. Энкодер преобразует входные данные в компактное скрытое

Автокодировщики — продолжение

Для генерации новых объектов необходимо, чтобы скрытое пространство было организовано так, чтобы из него можно было корректно выбирать новые точки.

Сравнительная таблица: Автокодировщики

Автокодировщики	Автокодировщик состоит из двух основных частей: энкодера и декодера. Энкодер преобразует входные данные в компактно
Вариационные автокодировщики VAE	Обычный автокодировщик прежде всего решает задачу реконструкции. Для генерации новых объектов необходимо, чтобы скр

Вариационные автокодировщики VAE

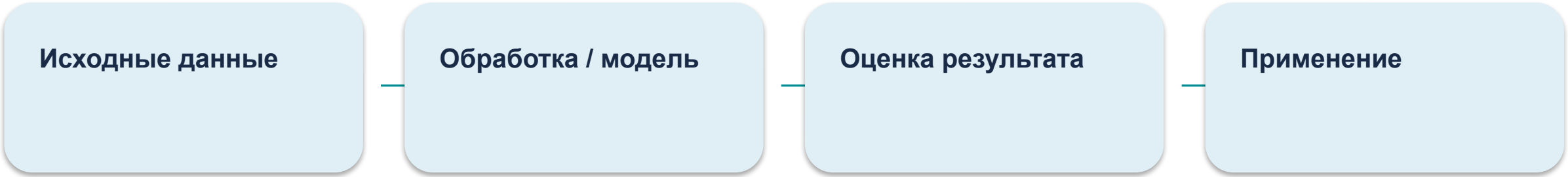
Variational Autoencoder (VAE) — генеративная модификация автокодировщика. Вместо фиксированной точки энкодер описывает вероятностное распределение в скрытом пространстве. При обучении модель одновременно стремится хорошо реконструировать данные и поддерживать упорядоченную структуру скрытого пространства.

Ключевая идея: Variational Autoencoder (VAE) — генеративная модификация автокодировщика. Вместо фиксированной точки энкодер описывает вероят

Вариационные автокодировщики VAE — продолжение

После обучения из скрытого распределения можно выбрать новую точку и передать её декодеру, получив новый объект. В загруженных учебных материалах отдельно предусмотрено изучение автокодировщиков, VAE и CVAE как современных нейросетевых подходов.

Схема: Вариационные автокодировщики VAE



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Условные генеративные модели

Условная генеративная модель создаёт данные не полностью свободно, а с учётом дополнительного условия. Условием может быть класс, текстовое описание, изображение, стиль, категория или иной контекст. Например, conditional VAE может генерировать объект заданного класса, а модель text-to-image создаёт изображение по текстовому описанию.

Ключевая идея: Условная генеративная модель создаёт данные не полностью свободно, а с учётом дополнительного условия. Условием может быть кл

Условные генеративные модели — продолжение

Управляемость генерации является одним из ключевых свойств современных генеративных систем.

Генеративно-состязательные сети GAN

Generative Adversarial Network (GAN) включает две нейронные сети: генератор и дискриминатор. Генератор создаёт синтетические примеры, а дискриминатор пытается отличить реальные данные от созданных генератором.

Ключевая идея: Generative Adversarial Network (GAN) включает две нейронные сети: генератор и дискриминатор. Генератор создаёт синтетические

Генеративно-состязательные сети GAN — продолжение

Рисунок 2 – Упрощённая схема генеративно-состязательной сети В ходе обучения генератор стремится создавать всё более правдоподобные примеры, а дискриминатор — лучше различать реальные и синтетические данные. Такой состязательный процесс может приводить к высокому качеству генерации, особенно для изображений.

Диффузионные модели

Диффузионные модели стали важным классом генеративных моделей для изображений, аудио и видео. Идея основана на последовательном зашумлении данных и обучении модели обратному процессу — постепенному восстановлению структуры из шума. Рисунок 3 – Концептуальная схема диффузионной генерации Во время генерации процесс начинается с случайного шума.

Ключевая идея: Диффузионные модели стали важным классом генеративных моделей для изображений, аудио и видео. Идея основана на последовательн

Диффузионные модели — продолжение

Модель шаг за шагом удаляет шум, направляя результат к объекту, согласованному с условием. При text-to-image таким условием является текстовое описание.

Диффузионные модели — продолжение

В загруженной англоязычной хрестоматии приведён пример современной модели генерации видео, где видео сначала переводится в компактное скрытое пространство, представляется пространственно-временными патчами, а генерация выполняется диффузионным трансформером.

Схема: Диффузионные модели

Диффузионные модели стали важным классом генерат

Рисунок 3 – Концептуальная схема диффузионной ге

Во время генерации процесс начинается с случайно

В загруженной англоязычной хрестоматии приведён

Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Сравнительная таблица: Диффузионные модели

Диффузионные модели	Диффузионные модели стали важным классом генеративных моделей для изображений, аудио и видео. Идея основана на посл
GAN, VAE и диффузионные модели: сравнение	Рисунок 3 – Концептуальная схема диффузионной генерации
Генерация текста	Во время генерации процесс начинается с случайного шума. Модель шаг за шагом удаляет шум, направляя результат к объ
Генерация программного кода	В загруженной англоязычной хрестоматии приведён пример современной модели генерации видео, где видео сначала перево

Генерация текста

Генерация текста основана на моделировании последовательностей естественного языка. Современные языковые модели часто используют архитектуру Transformer и обучаются предсказывать токены или восстанавливать части текста в зависимости от постановки обучения.

Ключевая идея: Генерация текста основана на моделировании последовательностей естественного языка. Современные языковые модели часто использ

Генерация текста — продолжение

Генерируемый текст может использоваться для черновиков, объяснений, суммаризации, перевода, преобразования стиля, структурирования информации и программной помощи. Однако правдоподобность текста не означает его фактическую правильность. Например, модель может сформировать связное объяснение научного понятия, но допустить ошибку в числах, ссылках или фактах.

Генерация текста — продолжение

Поэтому результат необходимо проверять по надёжным источникам.

Генерация программного кода

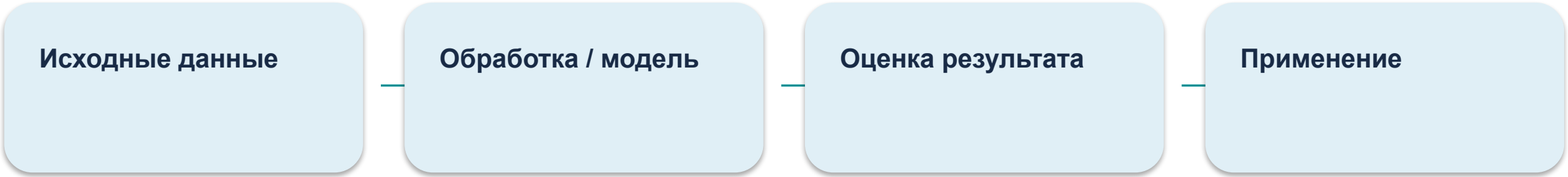
Программный код можно рассматривать как структурированную последовательность токенов. Генеративная модель способна создавать функции, объяснять код, предлагать тесты и преобразовывать программу между языками. При этом сгенерированный код необходимо проверять: возможны логические ошибки, неправильное использование библиотек, уязвимости или несоответствие требованиям.

Ключевая идея: Программный код можно рассматривать как структурированную последовательность токенов. Генеративная модель способна создавать

Генерация программного кода — продолжение

Генерация кода должна дополняться тестированием, анализом безопасности и контролем разработчика.

Схема: Генерация программного кода



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Генерация изображений

Генерация изображений может выполняться GAN, VAE и особенно диффузионными моделями. Пользователь задаёт текстовое условие, а модель формирует новое визуальное содержание. Современные системы часто работают не непосредственно с исходными пикселями на всех этапах, а используют компактное латентное пространство.

Ключевая идея: Генерация изображений может выполняться GAN, VAE и особенно диффузионными моделями. Пользователь задаёт текстовое условие, а

Генерация изображений — продолжение

Это уменьшает вычислительные затраты и позволяет эффективнее обучать и применять модель.

Сравнительная таблица: Генерация изображений

Генерация изображений	Генерация изображений может выполняться GAN, VAE и особенно диффузионными моделями. Пользователь задаёт текстовое у
Генерация аудио	Современные системы часто работают не непосредственно с исходными пикселями на всех этапах, а используют компактное

Генерация аудио

Генеративные модели могут создавать речь, музыку и другие звуковые сигналы. Задача может быть условной: текст преобразуется в речь, голосовой сигнал модифицируется, а музыкальный фрагмент создаётся по описанию или заданной структуре. Аудиогенерация требует особого внимания к правам на голос, персональным данным и возможности введения слушателя в заблуждение.

Ключевая идея: Генеративные модели могут создавать речь, музыку и другие звуковые сигналы. Задача может быть условной: текст преобразуется в

Генерация аудио — продолжение

Синтетический голос должен использоваться с разрешением и в соответствии с правовыми и этическими требованиями.

Генерация видео

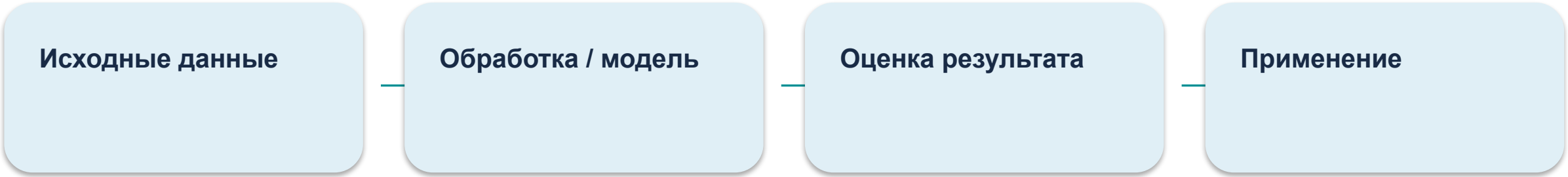
Генерация видео сложнее генерации отдельного изображения, поскольку необходимо поддерживать согласованность не только внутри кадра, но и во времени. Модель должна учитывать движение объектов, изменение камеры, освещения и причинно-временные зависимости между кадрами.

Ключевая идея: Генерация видео сложнее генерации отдельного изображения, поскольку необходимо поддерживать согласованность не только внутри

Генерация видео — продолжение

В англоязычной хрестоматии рассмотрен подход, в котором видео сжимается в латентное пространство, делится на пространственно-временные патчи, а диффузионный трансформер обучается восстанавливать чистое представление из зашумленного. Такой подход демонстрирует объединение идей компрессии, Transformer и diffusion.

Схема: Генерация видео



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Мультимодальный генеративный ИИ

Мультимодальная модель работает с несколькими типами данных. Например, она может получать текст и изображение, анализировать их совместно и формировать текстовый ответ; либо получать текст и создавать изображение или видео.

Prompt как условие генерации

Prompt — инструкция или контекст, передаваемый модели для управления результатом. Формулировка запроса влияет на содержание, формат и уровень детализации ответа. Для текстовой модели полезно явно задавать задачу, входные данные, ограничения и ожидаемый формат. Для генерации изображений описание может включать объект, композицию, освещение, стиль и другие параметры.

Ключевая идея: Prompt — инструкция или контекст, передаваемый модели для управления результатом. Формулировка запроса влияет на содержание,

Prompt как условие генерации — продолжение

Более системно методы построения запросов рассматриваются в отдельной теме Prompt Engineering.

Сравнительная таблица: Prompt как условие генерации

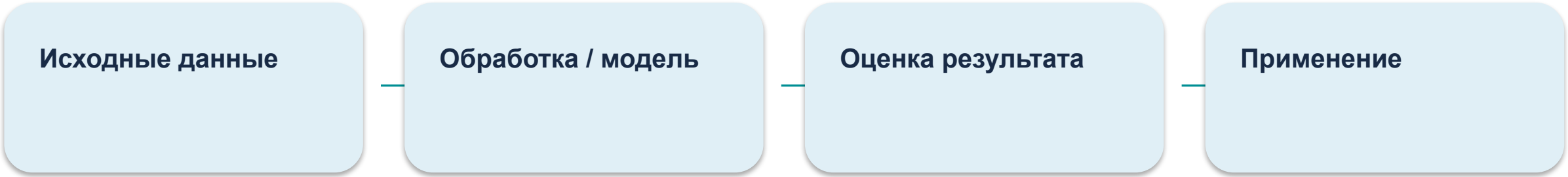
Prompt как условие генерации	Prompt — инструкция или контекст, передаваемый модели для управления результатом. Формулировка запроса влияет на со
Параметры разнообразия генерации	Для текстовой модели полезно явно задавать задачу, входные данные, ограничения и ожидаемый формат. Для генерации из

Параметры разнообразия генерации

Генеративная модель часто формирует вероятностное распределение возможных результатов. Управление выбором из этого распределения позволяет менять степень вариативности. Параметры генерации нельзя оценивать как «лучше» или «хуже» вне задачи. Для формального технического текста обычно требуется большая стабильность, а для творческого поиска может быть полезно разнообразие.

Ключевая идея: Генеративная модель часто формирует вероятностное распределение возможных результатов. Управление выбором из этого распределе

Схема: Параметры разнообразия генерации



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Оценка качества генеративного результата

Оценка генеративных моделей сложнее, чем оценка обычного классификатора, поскольку часто не существует единственного правильного результата. Необходимо учитывать несколько критериев одновременно.

Галлюцинации и фактическая достоверность

Для текстовых генеративных моделей характерна возможность сформировать правдоподобное, но фактически неверное утверждение. Такое явление часто называют галлюцинацией модели. Причина заключается в том, что модель оптимизируется для построения вероятного продолжения или ответа в рамках обученного представления, а не гарантированного обращения к проверенной базе фактов.

Ключевая идея: Для текстовых генеративных моделей характерна возможность сформировать правдоподобное, но фактически неверное утверждение. Та

Галлюцинации и фактическая достоверность — продолжение

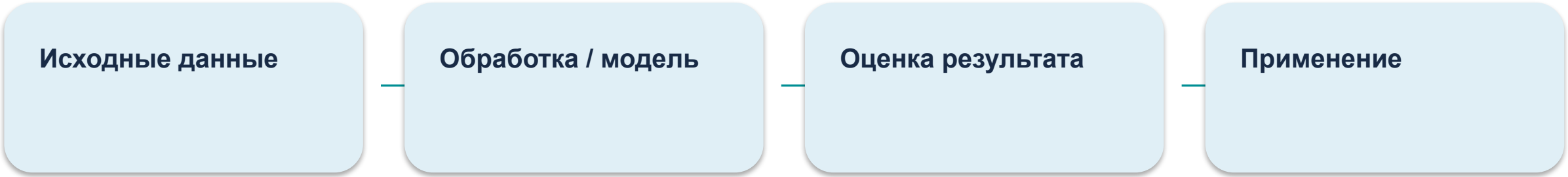
Поэтому фактические сведения, числа, ссылки, нормативные положения и научные утверждения необходимо проверять по первичным или надёжным источникам.

Смещения и качество обучающих данных

Генеративная модель наследует свойства данных, на которых она обучалась. Если данные содержат ошибки, дисбаланс, устаревшие сведения или систематические искажения, это может проявляться в результатах. Ответственное применение требует учитывать происхождение данных, область применимости модели и возможные различия качества для разных языков, доменов и категорий пользователей.

Ключевая идея: Генеративная модель наследует свойства данных, на которых она обучалась. Если данные содержат ошибки, дисбаланс, устаревшие с

Схема: Смещения и качество обучающих данных



Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Сравнительная таблица: Смещения и качество обучающих данных

Смещения и качество обучающих данных	Генеративная модель наследует свойства данных, на которых она обучалась. Если данные содержат ошибки, дисбаланс, ус
Авторское право и происхождение контента	Ответственное применение требует учитывать происхождение данных, область применимости модели и возможные различия к

Авторское право и происхождение контента

Генеративный ИИ поднимает вопросы прав на обучающие материалы, сгенерированный контент и использование результатов. Практическая политика зависит от законодательства, лицензий и условий конкретного сервиса.

Ключевая идея: Генеративный ИИ поднимает вопросы прав на обучающие материалы, сгенерированный контент и использование результатов. Практичес

Авторское право и происхождение контента — продолжение

В учебной и научной работе необходимо прозрачно указывать использование генеративных инструментов, если этого требуют правила организации или издания, и не представлять автоматически созданный материал как проверенный первичный источник.

Персональные и конфиденциальные данные

Нельзя без необходимости передавать внешнему генеративному сервису персональные данные, пароли, ключи доступа, внутренние документы, закрытый исходный код и другую конфиденциальную информацию. Если генеративный ИИ используется в организации, необходимо определить разрешённые типы данных, права доступа, правила хранения запросов и результатов, а также процедуры проверки и публикации.

Ключевая идея: Нельзя без необходимости передавать внешнему генеративному сервису персональные данные, пароли, ключи доступа, внутренние док

Риски злоупотребления генеративными технологиями

Генеративные модели могут использоваться для создания вводящего в заблуждение контента, масштабирования социальной инженерии и подделки медиаматериалов. Поэтому средства генерации должны применяться с учётом безопасности и ответственности.

Ключевая идея: Генеративные модели могут использоваться для создания вводящего в заблуждение контента, масштабирования социальной инженерии

Риски злоупотребления генеративными технологиями — продолжение

Защитные меры включают обучение пользователей, проверку источников, многофакторную аутентификацию, контроль публикаций и использование технических средств проверки контекста и происхождения данных. Нельзя полагаться только на визуальную или стилистическую правдоподобность материала.

Сравнительная таблица: Риски злоупотребления генеративными технологиями

Риски злоупотребления генеративными технологиями	Генеративные модели могут использоваться для создания вводящего в заблуждение контента, масштабирования социальной
Генеративный ИИ в образовании и науке	Защитные меры включают обучение пользователей, проверку источников, многофакторную аутентификацию, контроль публика

Генеративный ИИ в образовании и науке

Генеративные модели могут помогать объяснять материал, создавать черновые примеры, формировать вопросы, структурировать идеи и поддерживать программирование. Однако они не должны подменять понимание предмета и самостоятельную проверку.

Ключевая идея: Генеративные модели могут помогать объяснять материал, создавать черновые примеры, формировать вопросы, структурировать идеи

Генеративный ИИ в образовании и науке — продолжение

В научной работе модель может помочь отредактировать формулировку, структурировать текст или предложить код анализа, но научные выводы, ссылки, экспериментальные результаты и интерпретация должны подтверждаться исследователем.

Генеративный ИИ в работе с документами

Практический сценарий: пользователь имеет длинный документ и хочет получить краткое содержание. Безопасная схема включает передачу только разрешённого содержания, формулирование задачи, получение чернового резюме, сравнение с исходным документом и ручное исправление.

Ключевая идея: Практический сценарий: пользователь имеет длинный документ и хочет получить краткое содержание. Безопасная схема включает пер

Генеративный ИИ в работе с документами — продолжение

Рисунок 4 – Рекомендуемый рабочий процесс при использовании генеративного ИИ Такой подход подчёркивает, что модель является инструментом подготовки результата, а ответственность за итоговое решение остаётся у человека или организации, применяющей систему. Пример сравнения генеративных подходов

Схема: Генеративный ИИ в работе с документами

Практический сценарий: пользователь имеет длинны

Рисунок 4 – Рекомендуемый рабочий процесс при ис

Такой подход подчёркивает, что модель является и

Пример сравнения генеративных подходов

Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Ограничения генеративных моделей

Результат может быть фактически неверным, хотя выглядит убедительно. Качество зависит от обучающих данных и области применения модели. Сложно гарантировать полное отсутствие смещений и нежелательных результатов. Некоторые модели требуют значительных вычислительных ресурсов. Оценка качества часто включает человеческое экспертное суждение.

Ключевая идея: Результат может быть фактически неверным, хотя выглядит убедительно.

Ограничения генеративных моделей — продолжение

Возникают вопросы авторского права, конфиденциальности и ответственности. Модель не должна автоматически получать доступ к критичным действиям без контроля и ограничений.

Основные выводы

Генеративный искусственный интеллект создаёт новый контент на основе закономерностей обучающих данных. К основным генеративным подходам относятся авторегрессионные модели, VAE, GAN и диффузионные модели. Современные системы могут генерировать текст, изображения, код, аудио и видео.

Ключевая идея: Генеративный искусственный интеллект создаёт новый контент на основе закономерностей обучающих данных.

Основные выводы — продолжение

Условная генерация позволяет управлять результатом с помощью текста, класса, изображения или другого контекста. Качество генерации необходимо оценивать по соответствию задаче, корректности, разнообразию и безопасности. Правдоподобный ответ генеративной модели может быть неверным и требует проверки.

Основные выводы — продолжение

При работе с генеративным ИИ необходимо учитывать конфиденциальность, авторское право, безопасность и ответственность.

Схема: Основные выводы

Генеративный искусственный интеллект создаёт нов

К основным генеративным подходам относятся автор

Современные системы могут генерировать текст, из

Условная генерация позволяет управлять результат

Последовательность отражает основные этапы рассматриваемого процесса; конкретная реализация зависит от задачи.

Контрольные вопросы

1. Что такое генеративный искусственный интеллект и чем генеративная модель отличается от дискриминативной? 2. Каковы основные принципы авторегрессионной генерации, VAE, GAN и диффузионных моделей? 3. В чём различие между обычным автокодировщиком и вариационным автокодировщиком? 4. Как взаимодействуют генератор и дискриминатор в GAN и какие проблемы могут возникать при обучении? 5.

Ключевая идея: 1. Что такое генеративный искусственный интеллект и чем генеративная модель отличается от дискриминативной?

Контрольные вопросы — продолжение

Как устроен общий принцип диффузионной генерации и почему он применяется для изображений и видео? 6. Какие особенности имеют генерация текста, изображений, аудио, видео и программного кода? 7. Что такое условная и мультимодальная генерация и какую роль в них играет prompt? 8. По каким критериям следует оценивать качество генеративного результата? 9.

Контрольные вопросы — продолжение

Какие риски связаны с галлюцинациями, конфиденциальными данными, авторским правом и вводящим в заблуждение контентом? 10. Как организовать безопасный и ответственный рабочий процесс применения генеративного ИИ в учебной, научной или профессиональной задаче?

Список источников

1. Учебное пособие «Основы искусственного интеллекта». Глава 11 «Генеративный искусственный интеллект»: Generative AI, генерация текста, изображений, аудио и видео. 2. Беликова Е.К., Гудилина И.С., Круглова М.А., Саратовская Л.Б., Ямпольская А.В. English Reader in Artificial Intelligence = Хрестоматия по искусственному интеллекту на английском языке. Учебное пособие. Москва, 2024.

Ключевая идея: 1. Учебное пособие «Основы искусственного интеллекта». Глава 11 «Генеративный искусственный интеллект»: Generative AI, генера

Список источников — продолжение

– 171 с. 3. Бессмертный И.А., Авдюшина А.Е., Кугаевских А.В., Королева Ю.А., Рущенко Н.Г. Системы искусственного интеллекта. Методическое пособие. Санкт-Петербург: Университет ИТМО, 2024. – 281 с. 4. Андреянов Н.В., Евдокимова Т.С., Павлов А.Д., и др. Методы искусственного интеллекта. Учебное пособие. Казань: КНИТУ-КАИ, 2024. – 392 с.

График: иллюстрация зависимости



Иллюстративные значения (не результаты эксперимента).

Итоги и вопросы для обсуждения

1. Объясните понятие «Понятие генеративного искусственного интеллекта» и приведите пример.
2. Объясните понятие «Генеративные и дискриминативные модели» и приведите пример.
3. Объясните понятие «Что значит «генерировать» данные» и приведите пример.
4. Объясните понятие «Основные классы генеративных моделей» и приведите пример.
5. Объясните понятие «Авторегрессионная генерация» и приведите пример.
6. Объясните понятие «Автокодировщики» и приведите пример.
7. Объясните понятие «Вариационные автокодировщики VAE» и приведите пример.