СӘТБАЕВ УНИВЕРСИТЕТІ SATBAYEV UNIVERSITY

Основы искусственного интеллекта. Лекция 8. Генеративный искусственный интеллект и большие языковые модели (LLM)

Преподаватель: **Қасенхан Арай Мейрамбайқызы, доктор PhD, ассоциированный профессор кафедры «Программная инженерия»**

a.kassenkhan@satbayev.university

Содержание

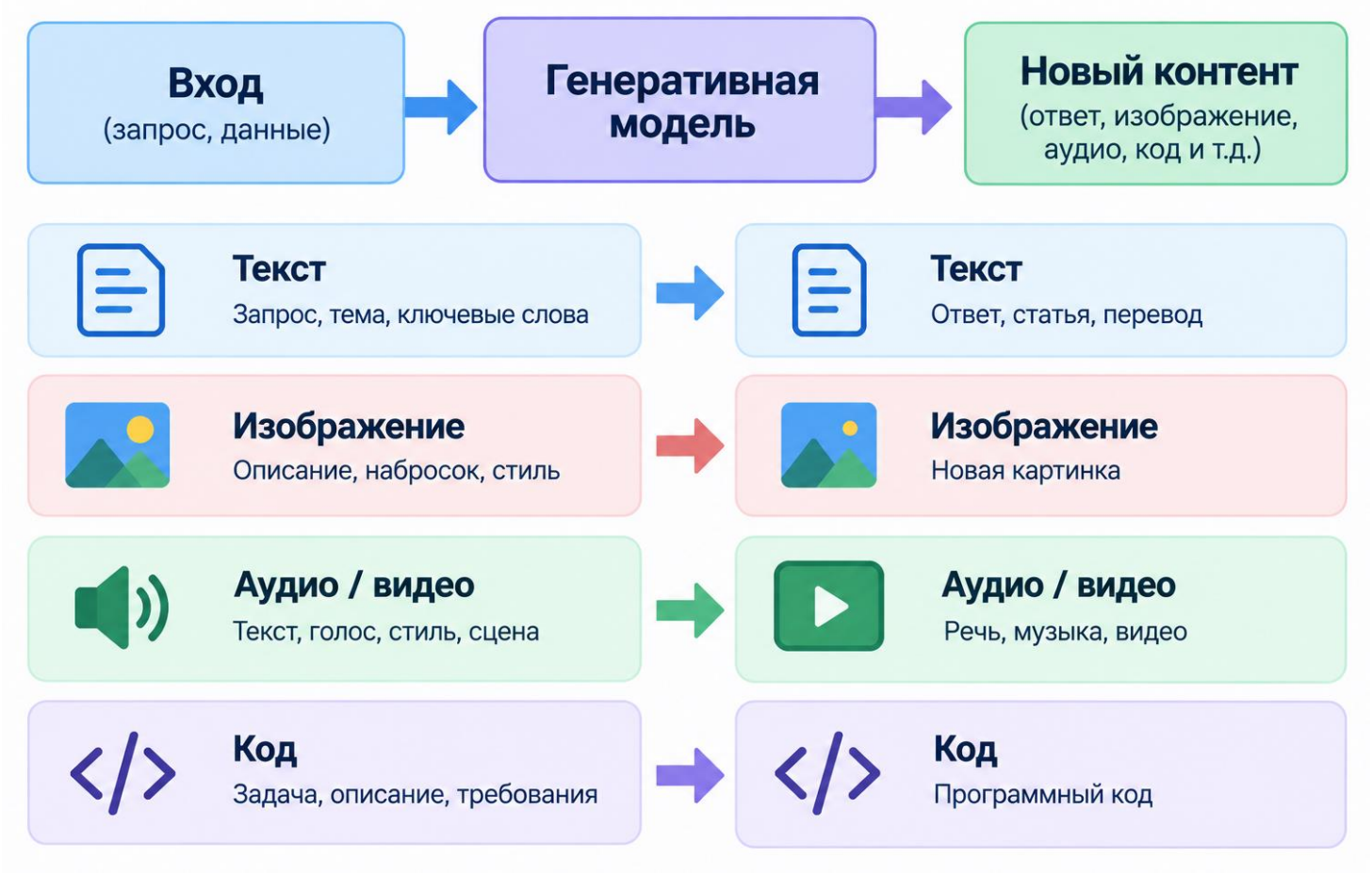
1. Что такое генеративный ИИ
2. Как работают большие языковые модели
3. Transformer и генерация текста
4. Prompt Engineering
5. Возможности и ограничения генеративного ИИ

По завершению урока Вы будете знать:

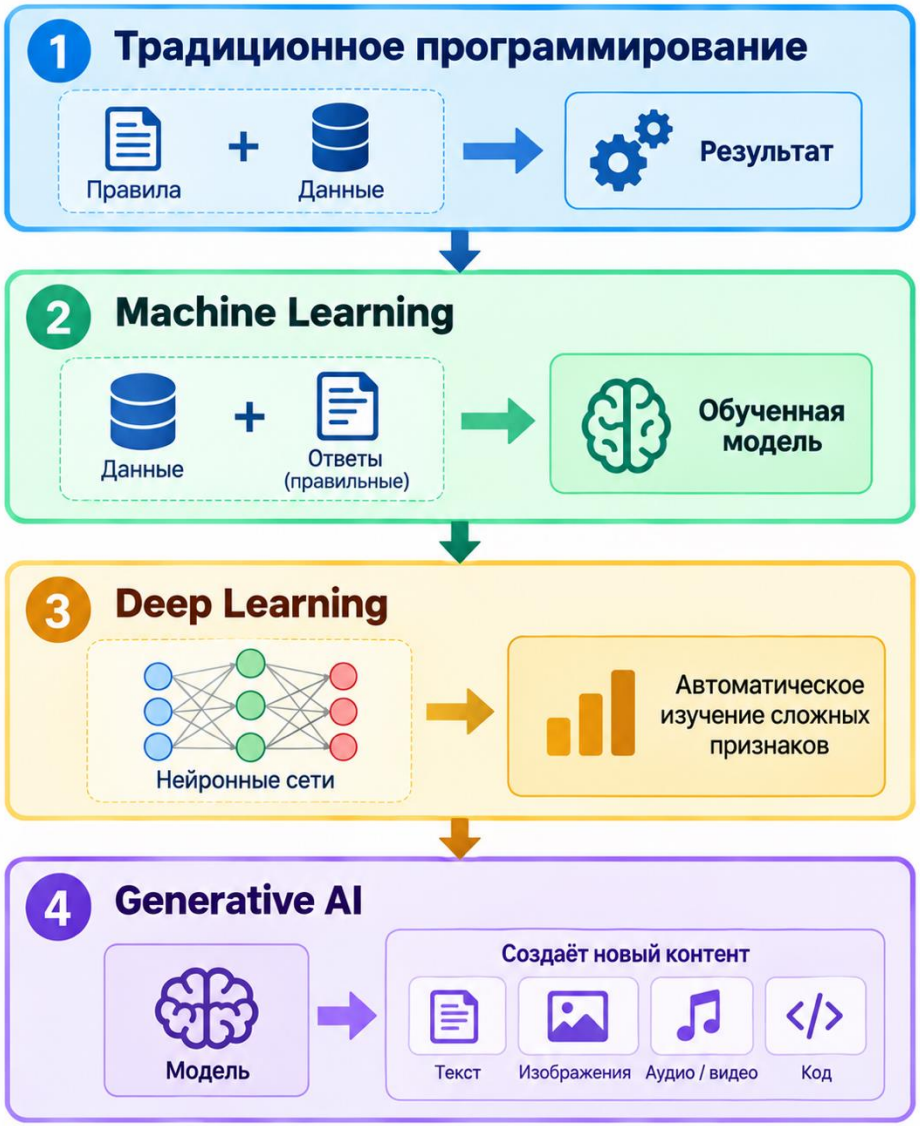
1. **Чем генеративный ИИ отличается от традиционного ИИ**
2. **Что такое LLM и как она обучается**
3. **Зачем нужны токены, embeddings и механизм attention**
4. **Как LLM генерирует текст**
5. **Как составлять эффективные промпты**
6. **Почему LLM может ошибаться и как проверять её ответы**

Что такое генеративный ИИ?

Генеративный ИИ создаёт новый контент на основе закономерностей, изученных из данных.

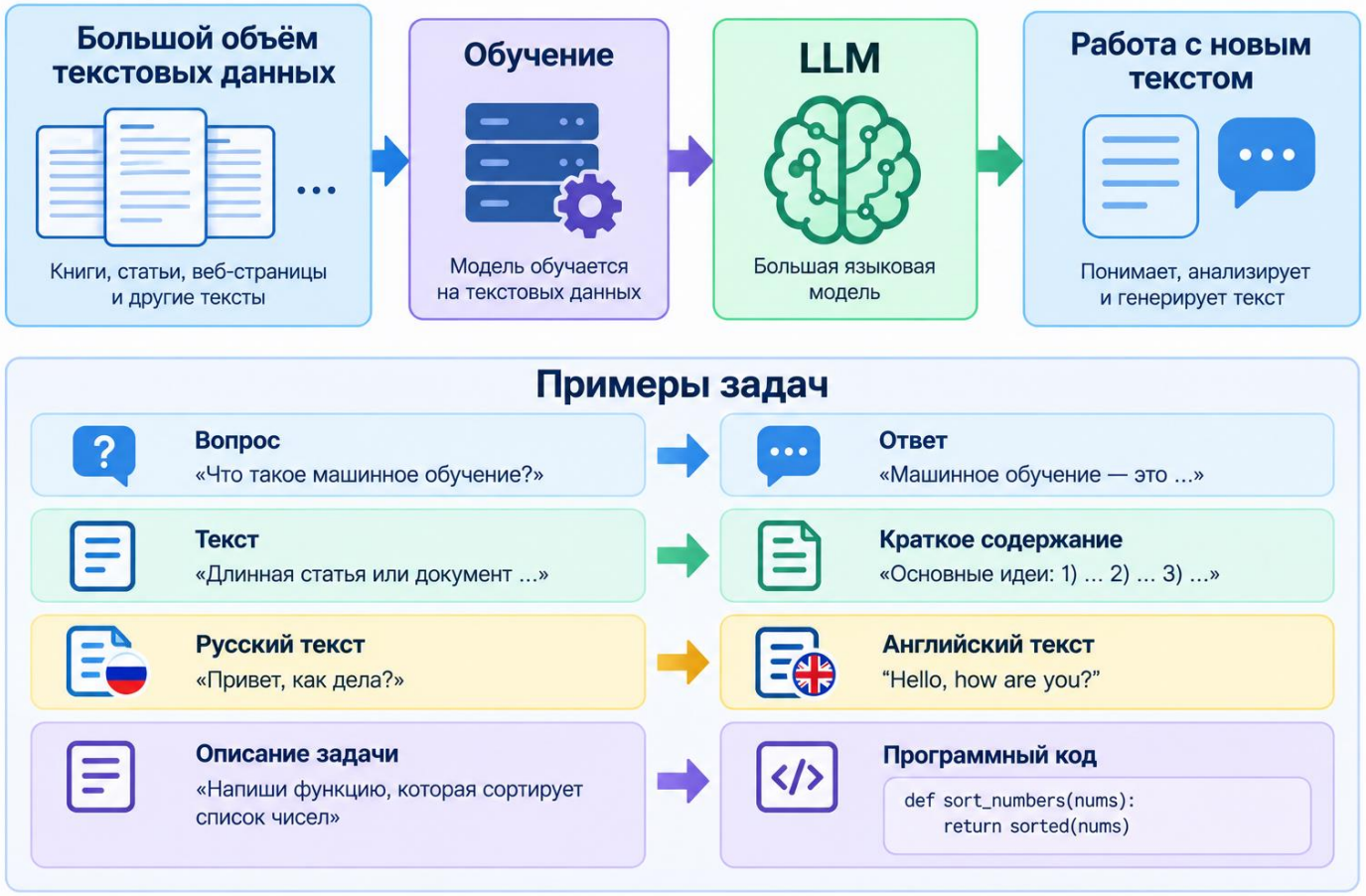


От машинного обучения к генеративному ИИ



Что такое LLM?

LLM — большая языковая модель, обученная работать с закономерностями языка.



Как обучается LLM?

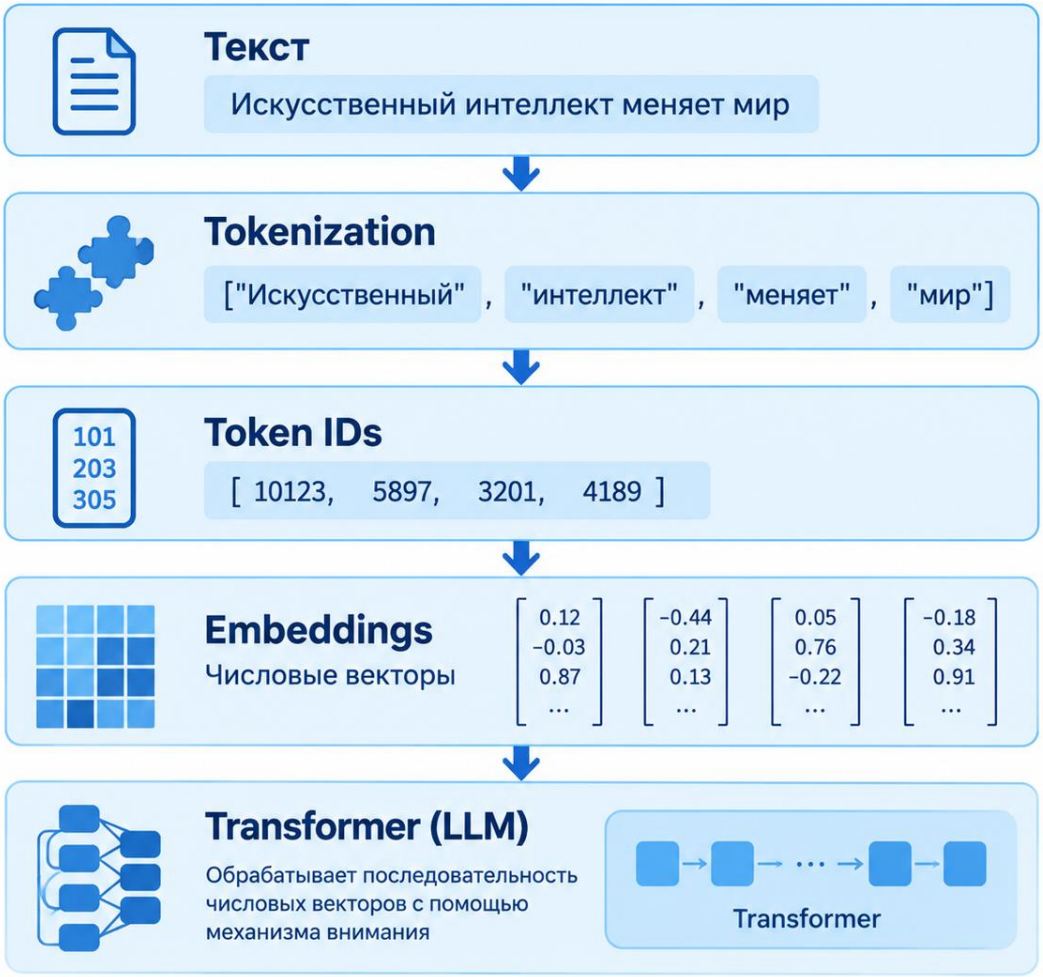
1. Сбор данных
2. Pre-training
3. Fine-tuning
4. Alignment

Pre-training → модель изучает закономерности языка

Fine-tuning → адаптация под задачи

Alignment → настройка поведения модели

Как текст превращается в числа?



LLM получает не слова напрямую, а их числовое представление.

Transformer: ключевая идея



Attention помогает модели учитывать связь между словами в контексте.

Как LLM генерирует текст?



LLM генерирует ответ последовательно, предсказывая следующий токен на основе предыдущего контекста.

Что такое Prompt?

Слабый prompt

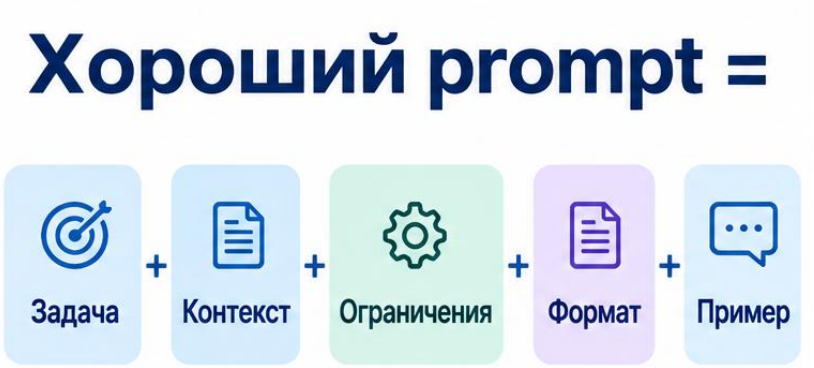
Расскажи про искусственный интеллект.

Более точный prompt

Объясни студенту 1 курса, что такое искусственный интеллект. Используй простой язык, приведи 3 примера и ограничь ответ 150 словами.

Чем точнее задача и контекст, тем более управляемым становится результат.

Структура хорошего промпта



Чем более конкретный, полный и структурированный prompt, тем лучше результат.



Zero-shot, One-shot и Few-shot

Один и тот же запрос, но разное количество примеров в промпте.

Zero-shot

Без примера

Промпт:

Определи тональность:
"Мне очень понравился курс".

Ответ модели:

Позитивный

One-shot

Один пример

Промпт:

"Курс был скучным" → негативный

"Мне очень понравился курс" → ?

Ответ модели:

Позитивный

Few-shot

Несколько примеров

Промпт:

"Курс был скучным" → негативный
"Материал интересный" → позитивный
"Сложно было разобраться" → негативный
"Мне очень понравился курс" → ?

Ответ модели:

Позитивный

Zero-shot
→ без примера

Модель выполняет задачу без примеров.

One-shot
→ один пример

Модель получает один пример, чтобы понять формат и выполнить задачу.

Few-shot
→ несколько примеров

Модель получает несколько примеров, чтобы лучше понять закономерность и выполнить задачу.

LLM может ошибаться

1 Галлюцинации

Модель может уверенно сообщить неверную информацию.



Первый президент Казахстана — Нурсултан Назарбаев, а столица страны — Алматы.

✕ Неверная информация

2 Предвзятость

Ответ зависит от данных и контекста обучения.



Например, модель может воспроизводить стереотипы или давать несправедливые ответы.

3 Конфиденциальность

Нельзя бездумно передавать чувствительные данные.



Не передавайте:

- личные данные (ФИО, адрес, ИИН)
- корпоративную информацию
- медицинские данные
- финансовые данные

4 Авторское право

Сгенерированный контент требует ответственного использования.



Важно:

- проверять источники
- не выдавать сгенерированный контент за свой
- соблюдать лицензионные требования

Ответ LLM ≠ гарантированный факт

Проверяйте важную информацию по надёжным источникам!!!

Заключение



Generative AI → создаёт новый контент
LLM → работает с языком и генерирует последовательности токенов
Transformer → учитывает контекст с помощью attention
Prompt Engineering → помогает управлять результатом

ИИ генерирует ответ, но ответственность за его проверку и использование остаётся за человеком.