

# Ответ 1

## Что такое векторное представление (embedding) в контексте искусственного интеллекта и какую основную задачу оно решает?

На основе предоставленного учебного материала, вот ответ на ваш вопрос.

### Основная задача: Поиск по смыслу

Векторные представления (embeddings) решают фундаментальную проблему: они позволяют компьютерам понимать не просто точные слова, а их **смысл**.

Традиционный поиск, например, на компьютере, работает по ключевым словам. Если вы ищете документ со словом «отчет», система найдет только те файлы, где есть именно это слово. Однако, если нужный документ называется «ежеквартальный итог» и не содержит слова «отчет», традиционный поиск его не обнаружит, хотя по смыслу он вам подходит.

Эта проблема решается с помощью **семантического поиска** — поиска по значению. Как указано в материале, основная задача, которую решают векторные представления, — это «представить значение слов, предложений, изображений и других данных в числовом формате», чтобы компьютер мог с ними работать и понимать их смысл. Это позволяет находить документы, которые *похожи по значению* на ваш запрос, а не только по точному совпадению слов.

### Что такое векторное представление (Embedding)?

В контексте искусственного интеллекта (**AI**), **векторное представление (embedding)** — это числовое представление сложных данных, таких как текст, изображения или аудио. По сути, это процесс преобразования этих данных в **вектор** — упорядоченный список чисел.

#### 1. От данных к числам

Как объясняется в учебном материале, в основе современных **LLM** (Large Language Models) лежит способность превращать данные в математические объекты. Таким объектом и является вектор.

**Простая аналогия из материала:** Представим, что мы описываем фильм тремя характеристиками: «Комедия», «Экшен» и «Драма» по шкале от 0 до 10.

- Фильм «Мстители» можно представить как вектор: [2, 9, 4] (мало комедии, много экшена, немного драмы).
- Романтическую комедию «Дневник Бриджит Джонс» — как вектор: [8, 1, 6].

Этот список чисел и есть вектор.

#### 2. Embeddings в современных AI-моделях

Современные AI-модели, такие как **GPT**, **Claude** или **Llama**, выполняют ту же операцию, но в гораздо большем масштабе. Они преобразуют фрагмент текста в вектор, состоящий из сотен или тысяч чисел (измерений). Этот процесс называется **созданием векторного представления** или **embedding**.

Полученный в результате вектор (embedding) является, как сказано в тексте, «числовым «слепок» или «ДНК» семантического значения исходных данных».

#### 3. Ключевой принцип работы

Главная идея векторных представлений заключается в следующем: **объекты с похожим смыслом будут иметь похожие векторы**.

Это означает, что в многомерном пространстве векторы для семантически близких понятий будут находиться рядом друг с другом. Материал приводит следующий пример:

- Векторы для предложений «Какая сегодня погода?» и «Будет ли сегодня дождь?» будут расположены очень близко.
- Вектор для предложения «Я люблю пиццу» будет находиться далеко от них.

Эти векторы создаются с помощью сложных нейронных сетей, таких как **трансформеры (transformers)**. В основе их

архитектуры лежит механизм **self-attention** (само-внимания), который позволяет модели оценивать важность разных слов в предложении для создания его общего векторного представления.