

# Рекомендации без обучения на основе таксономии с использованием больших языковых моделей

Дата: 2025-02-19 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2406.14043>

Рейтинг: 70

Адаптивность: 85

## Ключевые выводы:

Исследование направлено на решение проблем использования больших языковых моделей (LLM) в рекомендательных системах. Авторы предлагают новую структуру TAXREC, которая использует таксономию для категоризации элементов и улучшения рекомендаций LLM. Основной результат: TAXREC значительно превосходит традиционные подходы к рекомендациям с нулевым обучением, особенно в доменах, где LLM имеют ограниченные знания.

## Объяснение метода:

Исследование предлагает эффективный метод структурирования данных через таксономию для улучшения рекомендаций LLM. Основные концепции – двухэтапный подход и организация информации – могут быть адаптированы пользователями для различных задач, особенно при работе с большими объемами данных. Однако полная реализация требует технических навыков, что ограничивает прямую применимость для нетехнических пользователей.

**## Ключевые аспекты исследования:** 1. **Таксономия для структурирования данных:** Исследование предлагает подход TAXREC, который использует таксономию (систематическую классификацию) для организации и категоризации элементов (фильмов, книг и т.д.) перед их подачей в LLM для рекомендаций.

**Двухэтапный процесс:** Метод включает одноразовую классификацию элементов по таксономии (извлечение знаний из LLM о категориях и атрибутах) и последующую рекомендацию на основе LLM с использованием этой структурированной информации.

**Решение проблемы ограничения длины промпта:** TAXREC решает проблему ограниченной длины контекста LLM, сжимая большой пул элементов в компактную таксономию, что позволяет эффективно представлять информацию в пределах лимита токенов.

**Механизм сопоставления признаков:** Система преобразует неструктурированные выходные данные LLM в структурированный формат и использует механизм сопоставления для ранжирования рекомендаций на основе совпадения признаков.

**Улучшение рекомендаций в режиме zero-shot:** Исследование демонстрирует, что TAXREC значительно превосходит существующие методы рекомендаций в режиме zero-shot (без предварительного обучения на пользовательских данных).

## Дополнение: Для работы методов, описанных в исследовании, не требуется дообучение или API. Авторы используют стандартные модели (GPT-4 и Llama 2) через промпты, без дополнительного обучения. Все описанные подходы могут быть реализованы в стандартном чате с LLM.

Основные концепции и подходы, которые можно применить в стандартном чате:

**Извлечение таксономии из LLM:** Можно попросить модель создать систему категорий для любой области (книги, фильмы, товары и т.д.) с помощью промпта, подобного приведенному в Таблице 1 исследования.

**Структурирование элементов:** Можно запросить модель категоризировать отдельные элементы согласно созданной таксономии, обогащая их дополнительной контекстной информацией.

**Двухэтапный процесс запросов:** Сначала получить структурированное представление данных, затем использовать это представление для основной задачи (например, рекомендаций).

**Сжатие информации через таксономию:** Вместо перечисления всех элементов можно описать их категории и характеристики, что позволяет эффективно работать с большими наборами данных в пределах ограничений контекста.

**Структурирование выходных данных:** Можно запрашивать у модели ответы в определенном формате, основанном на таксономии, для облегчения последующей обработки.

Результаты применения этих концепций: - Более точные и релевантные ответы LLM благодаря лучшему пониманию контекста и структуры данных - Возможность работы с большими наборами данных в пределах ограничений контекста - Улучшенные рекомендации и ранжирование в режиме zero-shot (без дополнительного обучения) - Более структурированные и удобные для использования выходные данные

## Анализ практической применимости: 1. **Таксономия для структурирования данных** - Прямая применимость: Высокая. Пользователи могут применять принцип категоризации для улучшения запросов к LLM, организуя входные данные в структурированном виде. - Концептуальная ценность: Значительная. Понимание важности структурирования информации для LLM помогает пользователям

формулировать более эффективные запросы. - Потенциал для адаптации: Очень высокий. Подход можно адаптировать для любых задач, требующих организации большого количества элементов для LLM.

**Двухэтапный процесс** Прямая применимость: Средняя. Рядовым пользователям может быть сложно самостоятельно реализовать полный двухэтапный процесс. Концептуальная ценность: Высокая. Понимание преимуществ предварительной обработки данных перед основным запросом к LLM. Потенциал для адаптации: Высокий. Принцип разделения задачи на этапы категоризации и основного запроса может быть адаптирован для различных задач.

### **Решение проблемы ограничения длины промпта**

Прямая применимость: Высокая. Пользователи могут применять принцип сжатия информации через категоризацию для преодоления ограничений контекста LLM. Концептуальная ценность: Очень высокая. Понимание того, как структурирование и категоризация помогают преодолеть ограничения LLM. Потенциал для адаптации: Высокий. Подход применим для любых задач с большими объемами данных.

### **Механизм сопоставления признаков**

Прямая применимость: Средняя. Требуется технических навыков для реализации. Концептуальная ценность: Высокая. Демонстрирует важность структурирования выходных данных LLM для дальнейшей обработки. Потенциал для адаптации: Средний. Общий принцип сопоставления признаков может быть применен в упрощенном виде.

### **Улучшение рекомендаций в режиме zero-shot**

Прямая применимость: Высокая. Позволяет получать качественные рекомендации без обучения на пользовательских данных. Концептуальная ценность: Высокая. Демонстрирует потенциал LLM для задач рекомендаций без дополнительного обучения. Потенциал для адаптации: Высокий. Принципы zero-shot рекомендаций могут быть адаптированы для различных областей.

## **Prompt:**

Использование таксономии в промптах для рекомендательных систем на основе LLM ## Ключевые аспекты исследования TAXREC

Исследование показывает, что использование таксономии (структурированной категоризации) значительно улучшает качество рекомендаций, генерируемых большими языковыми моделями, особенно в областях с ограниченными знаниями модели.

## Пример промпта с использованием таксономии

[=====] Ты - рекомендательная система для книг. Я предоставляю тебе: 1. Историю

моих прочитанных книг с моими оценками 2. Таксономию для каждой книги (жанр, тематика, стиль написания, период, целевая аудитория)

Прочитанные книги: - "1984" Джордж Оруэлл (оценка: 5/5) Таксономия: {жанр: антиутопия, научная фантастика; тематика: тоталитаризм, контроль сознания; стиль: мрачный, философский; период: XX век; аудитория: взрослые} - "Гарри Поттер и философский камень" (оценка: 4/5) Таксономия: {жанр: фэнтези, приключения; тематика: взросление, дружба, магия; стиль: увлекательный; период: современный; аудитория: подростки, молодые взрослые}

На основе этой информации, пожалуйста: 1. Проанализируй мои предпочтения через призму таксономии 2. Рекомендуй 3 книги, которые могут мне понравиться 3. Для каждой рекомендации объясни, какие элементы таксономии повлияли на твой выбор [=====]

## Почему это работает

**Структурированное представление информации:** Таксономия помогает LLM лучше понимать характеристики элементов и связи между ними, преодолевая ограничения контекстного окна.

**Эффективное использование токенов:** Вместо передачи полной информации о множестве элементов, таксономия сжимает представление до ключевых характеристик.

**Улучшение рекомендаций в специализированных доменах:** Для областей, где у LLM ограниченные знания (например, специфические жанры книг), таксономия предоставляет структурированную информацию для более точных рекомендаций.

**Баланс детализации:** Исследование показывает, что оптимальное количество признаков в таксономии составляет 5-15 (слишком мало или слишком много снижает эффективность).

## Практическое применение

При создании промптов для рекомендательных систем: - Структурируйте информацию об элементах через таксономию - Включайте историю взаимодействий пользователя - Адаптируйте детализацию таксономии под конкретную задачу - Явно просите модель использовать таксономию для анализа предпочтений и генерации рекомендаций

Этот подход особенно эффективен, когда вы работаете с большими каталогами товаров или контента и хотите получить персонализированные рекомендации без обучения специализированной модели.