

# ТАРО: Адаптация, основанная на задаче, для оптимизации подсказок

Дата: 2025-02-26 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2501.06689>

Рейтинг: 65

Адаптивность: 75

## Ключевые выводы:

Исследование представляет ТАРО (Task-referenced Adaptation for Prompt Optimization) - фреймворк для оптимизации промптов, который динамически выбирает метрики оценки в зависимости от типа задачи и автоматизирует процесс эволюционной оптимизации промптов. Основным результатом - ТАРО превосходит существующие методы оптимизации промптов на различных задачах благодаря адаптивному подходу к выбору метрик и эволюционной оптимизации.

## Объяснение метода:

ТАРО предлагает ценные концепции адаптации промптов к типам задач, многокритериальной оценки и итеративного улучшения, применимые без технической реализации. Исследование демонстрирует эффективные стратегии для разных задач и моделей. Ограничения включают сложность полной реализации и необходимость API-доступа для некоторых компонентов.

**## Ключевые аспекты исследования:** 1. **Динамический выбор метрик:** ТАРО предлагает механизм динамического выбора метрик оценки для разных типов задач, адаптируя процесс оптимизации промптов к специфике конкретной задачи (например, точность для фактологических задач, разнообразие для творческих).

**Многокритериальная оценка промптов:** Система использует несколько метрик одновременно (сходство, разнообразие, сложность, перплексия) для комплексной оценки качества промптов, что позволяет более точно оптимизировать их для конкретных задач.

**Эволюционная оптимизация промптов:** Внедрение механизма эволюционной оптимизации через мутацию и отбор, что позволяет итеративно улучшать промпты, избегая локальных оптимумов.

**Адаптивность к разным типам задач:** ТАРО демонстрирует высокую адаптивность к различным типам задач (математические вычисления, рассуждения, перевод), что подтверждается экспериментами на шести разнотипных датасетах.

**Универсальность в отношении моделей:** Подход работает с различными LLM, включая проприетарные (GPT-3.5, GPT-4o) и открытые модели (Llama 3), показывая стабильное улучшение результатов по сравнению с базовыми методами.

## Дополнение: В исследовании TAPO используются внешние API и дообучение для экспериментальной валидации и количественной оценки результатов, однако ключевые концепции и подходы могут быть адаптированы для использования в стандартном чате без этих технических компонентов.

### Что можно применить в стандартном чате:

**Задачно-ориентированный промптинг:** Пользователи могут адаптировать свои запросы к типу задачи. Например: Для математических задач: "Разбей задачу на подзадачи, определи ключевые элементы, используй диаграммы или перефразируй, удали ненужную информацию" Для задач перевода: "Создай структурированную систему для категоризации ошибок, фокусируясь на распространённых проблемах и используй итеративное тестирование с обратной связью"

**Многокритериальная оценка:** Пользователи могут явно указывать несколько критериев в своих запросах:

"Ответ должен быть точным, но также разнообразным, с минимальными повторениями" "Необходимо обеспечить логическую последовательность и при этом сохранить простоту объяснения"

**Итеративное улучшение промптов:** Пользователи могут применять эволюционный подход вручную:

Начать с базового промпта Внести небольшие изменения в наиболее эффективные части Сохранить успешные модификации для будущего использования

**Стратегии из библиотеки TAPO:** Исследование предлагает конкретные стратегии, которые можно использовать в промптах:

"Разбей задачу на шаги" "Используй диаграммы для визуализации" "Удали ненужную информацию" ### Ожидаемые результаты:

Применение этих концепций в стандартном чате может привести к: - Более структурированным и понятным ответам LLM - Лучшей адаптации ответов к конкретному типу задачи - Повышению точности для фактологических задач и креативности для творческих - Более логичной последовательности рассуждений для сложных задач

Важно отметить, что хотя полная автоматизация процесса оптимизации промптов недоступна в стандартном чате, понимание и применение концептуальных основ TAPO может значительно улучшить качество взаимодействия с LLM.

**## Анализ практической применимости: 1. Динамический выбор метрик - Прямая применимость:** Средняя. Обычные пользователи не смогут напрямую реализовать сложную систему выбора метрик, но могут адаптировать идею выбора разных критериев оценки для разных типов задач. - **Концептуальная ценность:** Высокая. Понимание того, что разные задачи требуют разных критериев оценки ответов LLM, поможет пользователям формулировать более эффективные запросы. - **Потенциал для адаптации:** Высокий. Пользователи могут интуитивно адаптировать этот подход, указывая в промптах ожидаемые критерии качества ответа (точность, креативность, логическая последовательность).

**Многокритериальная оценка промптов Прямая применимость:** Низкая для технической реализации, но средняя для концептуального применения. **Концептуальная ценность:** Высокая. Понимание, что оценка по нескольким критериям дает лучшие результаты, поможет пользователям формулировать более комплексные запросы. **Потенциал для адаптации:** Средний. Пользователи могут интегрировать в свои запросы несколько критериев оценки, явно указывая, что ответ должен соответствовать нескольким параметрам.

### **Эволюционная оптимизация промптов**

**Прямая применимость:** Низкая в техническом смысле, но средняя в практическом применении. **Концептуальная ценность:** Высокая. Понимание процесса итеративного улучшения промптов может значительно повысить эффективность взаимодействия с LLM. **Потенциал для адаптации:** Высокий. Пользователи могут вручную применять итеративный подход к улучшению своих запросов, внося небольшие изменения и оценивая результаты.

### **Адаптивность к разным типам задач**

**Прямая применимость:** Средняя. Пользователи могут применять разные стратегии формулирования запросов для разных типов задач. **Концептуальная ценность:** Очень высокая. Понимание необходимости адаптации подхода к разным задачам критически важно для эффективной работы с LLM. **Потенциал для адаптации:** Очень высокий. Идея адаптации стратегий к типу задачи легко переносится на ручное формирование промптов.

### **Универсальность в отношении моделей**

**Прямая применимость:** Высокая. Результаты показывают, что подходы работают с разными моделями. **Концептуальная ценность:** Средняя. Понимание, что эффективные стратегии промптинга работают с разными моделями, полезно, но не трансформирует взаимодействие. **Потенциал для адаптации:** Высокий. Пользователи могут применять одни и те же стратегии к разным моделям с ожиданием похожих улучшений.

### **Prompt:**

Применение методологии ТАРО в промптах для GPT ## Ключевые принципы ТАРО для улучшения промптов

ТАРО (Task-referenced Adaptation for Prompt Optimization) позволяет оптимизировать промпты с помощью: - Выбора метрик оценки в зависимости от типа задачи - Многокритериальной оценки качества промптов - Эволюционной оптимизации для улучшения результатов

## Пример применения ТАРО для математической задачи

### Обычный промпт: [=====] Реши задачу: У Анны было 24 яблока. Она отдала  $\frac{1}{3}$  своих яблок Пете и  $\frac{1}{4}$  оставшихся яблок Марии. Сколько яблок осталось у Анны? [=====]

### Оптимизированный промпт с применением принципов ТАРО: [=====] # Математическая задача: распределение яблок

## Контекст У Анны было 24 яблока. Она отдала часть яблок двум друзьям.

## Шаги решения 1. Сначала выпиши начальное количество яблок у Анны 2. Рассчитай, сколько яблок Анна отдала Пете ( $\frac{1}{3}$  от начального количества) 3. Определи, сколько яблок осталось у Анны после передачи Пете 4. Рассчитай, сколько яблок Анна отдала Марии ( $\frac{1}{4}$  от оставшегося количества) 5. Вычисли финальное количество яблок у Анны

## Важно - Разбей решение на отдельные этапы - Проверь каждый шаг вычислений - Визуализируй процесс распределения яблок - Запиши ответ в формате "У Анны осталось X яблок" [=====]

## Как работает оптимизация по методологии ТАРО

**Адаптация к типу задачи:** Для математической задачи используется структурированный подход с разбиением на шаги, что повышает точность рассуждений.

**Многокритериальная оценка:** Промпт учитывает несколько критериев качества:

Схожесть (соответствие задаче) Логическую согласованность (пошаговое решение) Ясность (четкая структура)

**Эволюционная оптимизация:** Промпт включает элементы, которые показали высокую эффективность:

Разбиение на шаги Визуализация Проверка результатов Четкие инструкции по форматированию ответа Такой подход, согласно исследованию ТАРО, может повысить точность решения математических задач примерно на 10-15% по сравнению с базовыми промптами, особенно при использовании моделей GPT-4o и GPT-3.5-turbo.

