

Санкт-Петербургский политехнический университет
Петра Великого
Физико-механический институт
Высшая школа прикладной математики и вычислительной физики

Отчет по лабораторной работе №4
по дисциплине
"Автоматизация научных исследований"

Анализ промптов и моделей для структурного анализа научного текста

Выполнил студент:
Бабахина Софья Александровна
Группа: 5040102/50201

Санкт-Петербург
2025

Содержание

1 Цель работы	2
2 Задание	2
3 Результаты	4
3.1 Промпт 1	4
3.2 Промпт 2	4
3.3 Промпт 3	4
3.4 Промпт 4	4
3.5 Промпт 5	5
4 Выводы	5

1 Цель работы

Провести сравнительный анализ эффективности четырех заданных промптов и двух выбранных моделей ИИ для выявления структурных элементов и критической оценки научной статьи. На основе полученного анализа разработать оптимизированный авторский промпт, обеспечивающий максимально точный, полный и структурированный результат.

2 Задание

В рамках лабораторной работы необходимо произвести:

1. **Выбор и подготовка статьи.** Найти одну научную статью (желательно по теме, связанной с НИР или ВКР), в структуре которой заведомо присутствуют неявные или отсутствующие элементы.
 - **Рекомендация.** Для повышения сложности анализа моделям, необходимо подготовить текст статьи, удалив из него явные заголовки разделов («Введение», «Заключение» и т.п.).
2. **Выбор моделей.** Выбрать две различные ИИ-модели (например, Gemini, ChatGPT, GigaChat, DeepSeek и др.) для проведения эксперимента.
3. **Эксперимент.** Прогнать полный текст выбранной статьи через обе модели, используя каждый из четырех предложенных промптов.
4. **Анализ результатов.** Провести детальный сравнительный анализ 8 полученных ответов, выявляя сильные и слабые стороны различных моделей при работе с различными промптами.
5. **Оптимизация.** Выбрать одну ИИ-модель, которая, по вашему мнению, показала наилучший потенциал, и на основе выявленных недостатков существующих промптов, составить собственный промпт, который, по вашему мнению, должен дать наилучший результат по всем критериям.
6. **Финальная проверка.** Прогнать статью через выбранную модель, используя свой авторский промпт.

Критерии оценивания ответов ИИ

Критерии (по 10-балльной шкале) для оценки каждого из ответов:

Критерий	Описание
Полнота структурных элементов	Насколько полно ИИ выявил все элементы (актуальность, цель, задачи, методология, результаты, выводы).
Корректность и отсутствие выдуманной информации	ИИ не должен придумывать информацию, ошибочно определять методологию или путать выводы с результатами.
Критическая глубина анализа	Способность ИИ выявить логические противоречия, нарушения академического стиля, отсутствие связности цели и выводов.
Четкость формата ответа	Строгое соблюдение запрошенного в промпте формата (например, корректное отображение таблиц для Промта 3).

Выводы

В качестве выводов, в отчёте следует указать:

1. Сводная таблица оценок:

- Таблица с оценками (1–10 баллов) всех ответов.

2. Анализ сильных и слабых сторон:

- **Слабые стороны моделей:** Какие ошибки были характерны для каждой из двух выбранных моделей.
- **Слабые стороны промптов:** Какой из 4 промптов оказался наименее эффективным и почему.
- **Сильные стороны промптов:** Какой промпт оказался наиболее эффективным и почему.

3. Ваш собственный промпт:

- Полный текст разработанного промпта.
- Указание **модели**, для которой этот промпт разработан.

4. Обоснование и конечный результат:

- Подробное обоснование, **почему** авторский промпт был составлен именно так (какие проблемы предыдущих промпов он решает).
- Финальный ответ, полученный от ИИ с использованием разработанного промпта.
- Оценка этого финального ответа по 4 критериям и заключение об успехе оптимизации.

Промпты для нейросетей

Промпт 1

Проанализируй этот текст на соответствие структуре научной работы.

Промпт 2

Проанализируй научный текст на соответствие структуре. Найди, есть ли в нём актуальность, цели, задачи, предмет и объект исследования, методологию, результаты и выводы. Не придумывай информацию и опирайся только на текст.

Промпт 3

Проанализируй научный текст по структурным критериям научной статьи. Определи:

- актуальность,
- цель,
- задачи исследования,
- объект и предмет,
- методологию,
- результаты,

- выводы.

Представь ответ в виде таблицы:

Колонка 1 — элемент структуры.

Колонка 2 — найденный фрагмент текста (цитата).

Колонка 3 — краткий комментарий.

Если элемент отсутствует или выражен неявно — отметить это.

Промпт 4

Проанализируй научный текст с опорой на требования существующих ГОСТов и академические стандарты.

- проверь логическую связность между актуальностью, целью и задачами;
- выяви противоречия;
- оцени корректность формулировок;
- сравни цель и задачи (должны быть согласованы);
- отметить логические провалы;
- выдели отсутствующие элементы структуры;
- оцени соответствие выводов заявленным задачам.

Представь ответ структурировано: блок «оценка структуры», блок «ошибки», блок «вывод».

3 Результаты

3.1 Промпт 1

Ответ DeepSeek - файл 4_ds1.pdf .pdf

Ответ ChatGPT - файл 4_cg1.pdf .pdf

3.2 Промпт 2

Ответ DeepSeek - файл 4_ds2.pdf .pdf

Ответ ChatGPT - файл 4_cg2.pdf .pdf

3.3 Промпт 3

Ответ DeepSeek - файл 4_ds3.pdf .pdf

Ответ ChatGPT - файл 4_cg3.pdf .pdf

3.4 Промпт 4

Ответ DeepSeek - файл 4_ds4.pdf .pdf

Ответ ChatGPT - файл 4_cg4.pdf .pdf

3.5 Промпт 5

Текст собственного промпта:

Проанализируй предоставленный текст на соответствие структуре и стандартам научной работы (статьи или введения к исследованию). Действуй строго по следующему алгоритму:

1. Идентификация структурных элементов. Найди в тексте и представь в виде таблицы с колонками: «Элемент», «Цитата или констатация отсутствия», «Краткий комментарий».

Элементы для поиска: Актуальность исследования, Цель исследования, Задачи исследования, Объект и предмет исследования, Методология / методы, Основные результаты, Выводы

Если элемент отсутствует или выражен неявно - четко отметить это в соответствующей ячейке.

2. Глубинный анализ и оценка.

На основе выявленных элементов проведи критический разбор:

Логическая связность: Проследи, вытекают ли задачи из цели, а выводы - из результатов.

Оцени, решает ли работа заявленную цель.

Корректность формулировок: Проверь, являются ли цели и задачи конкретными, измеримыми, а предмет и объект - четко разграниченными.

Выявление противоречий и лакун: Отметь смысловые или логические противоречия между разделами, а также структурные пробелы (например, отсутствие методологии).

3. Итоговый вердикт.

Представь заключение в трех блоках:

Структурная полнота: Общая оценка наличия всех обязательных элементов научного текста.

Ключевые замечания и ошибки: Список основных проблем, выявленных на этапе глубинного анализа.

Общий вывод: Резюме о соответствии текста структуре научной работы с рекомендациями по устранению главных недостатков.

Важно: Всё время опирайся только на предоставленный текст, не придумывай и не домысливай информацию.

В качестве модели для проверки статьи по собственному промпту была выбрана модель DeepSeek. Результат представлен в файле 4_ds5pdf.pdf

4 Выводы

1. Таблица с оценками

Анализ сильных и слабых сторон:

- Слабые стороны моделей: один раз DeepSeek прогаллюцинировал, в остальном и чат и дипсик показали себя положительно (примерно одинаково)
- Слабые стороны промптов: все промпты выдают примерно одинаковые ответы, все отвечают корректно и подсвечивают все проблемные места в тексте правильно
- Сильные стороны промптов: наиболее эффективным оказался промпт 5, т.к. он наиболее структурирован и описывает четкую последовательность действий.

Заключение

- Промпты 1-3 обеспечивают стабильно хорошие, но базовые результаты. Они эффективны для простой идентификации элементов, но не направляют модель на глубокий анализ.
- Промпт 4, несмотря на высокие требования, в случае с DeepSeek привел к серьезной ошибке (галлюцинации), что подчеркивает риски излишне сложных или нечетко сформулированных запросов, даже для продвинутых моделей.

	Полнота структурных элементов	Корректность и отсутствие выдуманной информации	Критическая глубина анализа	Четкость формата ответа
Промпт 1				
DeepSeek	8	9	9	9
ChatGPT	9	9	9	9
Промпт 2				
DeepSeek	9	9	9	9
ChatGPT	9	9	9	9
Промпт 3				
DeepSeek 9	9	9	8	
ChatGPT	9	9	9	9
Промпт 4				
DeepSeek	9	1	9	9
ChatGPT (не совсем корректные данные, поскольку запрос был осуществлен через агрегатор ИИ Daisy)	8	10	8	8
Промпт 5				
DeepSekk	10	10	10	10

- Промпт 5 (синтезированный) доказал свою максимальную эффективность, получив высшие оценки по всем критериям. Его сила — в иерархической структуре: он последовательно направляет модель от простого извлечения данных (таблица) к их анализу и итоговой оценке, минимизируя риски ошибок и максимизируя полезность вывода для пользователя.

При работе с четкими инструкциями (Промпты 2, 3, 5) обе модели — DeepSeek и ChatGPT — показали практически идентичные и высокие результаты. Это указывает на то, что для стандартных задач анализа структуры текста выбор конкретной современной LLM-модели не является критичным. Единственный серьезный провал (оценка 1 по критерию корректности для DeepSeek на Промпте 4) служит важным напоминанием. Даже мощные модели могут «додумывать» информацию при работе с абстрактными или слишком сложными запросами. Наилучший промпт (№5) содержит явные страховки: требование опираться только на текст, указывать на отсутствие элементов и иметь четкую последовательность действий, что исключает пространные рассуждения «в пустоту». Оптимальный промпт должен сочетать структуру, запреты и последовательность. Исходя из результатов, наилучший промпт для анализа научного текста должен:

- Четко сегментировать задачу на логические этапы (выявление → анализ → вывод).
- Требовать формализованного вывода (таблица), что повышает точность и дисциплинирует модель.
- Содержать явные запреты на домыслы.
- Включать критерии глубинного анализа, но не как открытый вопрос, а как конкретный чек-лист для проверки (логическая связность, корректность формулировок).