

Санкт-Петербургский политехнический университет Петра Великого
Физико-механический институт

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 6
«Генерация списка литературы»
по дисциплине «Автоматизация научных исследований»

Выполнил
студент гр. № 5040102/50201

Зинкин С.В.

Преподаватель: Новиков Ф.А.

Санкт-Петербург
2025 г.

ЗАДАНИЕ

1. Процедура проведения

1. Для каждого из трех промптов (P1, P2, P3) задать идентичную задачу моделям Gemini и DeepSeek.
2. Из ответа каждой модели отобрать первые три уникальные и «живые» ссылки.
3. Для каждого из 18 отобранных источников провести верификацию и оценку по 5 описанным критериям.
4. Занести результаты в сводную таблицу (аналогичную приведенной) и рассчитать итоговые баллы.
5. Провести сравнительный анализ результатов, выявив сильные и слабые стороны каждой модели в зависимости от сложности промпта.

2. Ожидаемый результат (Анализ)

- Сводная таблица с оценками 18 источников.
- Выводы о сравнительной эффективности моделей Gemini и DeepSeek.
- Анализ влияния сложности промпта на качество результатов (например, рост баллов по критериям «Существование» и «Оформление» от P1 к P3).
- Оценка слабых мест моделей (например, недостаточное разнообразие типов источников).

3. Метрики оценки

Каждый предложенный моделью источник оценивается по критериям (максимум по каждому — 10 баллов):

1. Существование: Реальность источника.

- 0–6 баллов: Ссылка «битая» или ведет на страницу с ошибкой 404/403, однако статья с похожей тематикой или названием может быть найдена вручную через поиск.
- 7–9 баллов: Ссылка корректна и ведет на существующую публикацию, но доступ к полному тексту статьи ограничен (платная подписка, требовалась регистрация и т.д.).
- 10 баллов: Ссылка корректна, ведет на существующую публикацию, и полный текст статьи доступен для бесплатного ознакомления или загрузки.

2. Оформление: Соответствие ссылки требованиям ГОСТ.

Здесь включается форматирование той информации, которая была получена от AI. Проверка на соответствие госту. (0-10)

3. Полезность:

Оценивается релевантность найденного источника заявленной теме исследования (на основе анализа аннотации, ключевых слов, названия). 0–10 баллов: Субъективная оценка соответствия.

- 0 — тема полностью не совпадает,
- 10 — статья идеально соответствует запросу.

4. Новизна:

Оценивается актуальность источника на основе года его публикации.

- 0–4 балла: Год публикации до 1980.
- 5–7 баллов: Год публикации с 1981 по 1996.
- 7–8 баллов: Год публикации с 1997 по 2012.
- 9–10 баллов: Год публикации 2013 и новее.

5. Индекс цитируемости источника:

- 10 баллов: Q1.
- 8–9 баллов: Q2.
- 6–7 баллов: Q3.
- 5 баллов: Q4

- 1-4 балла: остальное

6. Разнообразие типа: Тип источника (статья, книга и т.д.) и его характеристики.

○ 0–8 баллов: Найденный источник — научная статья в журнале или сборнике.

○ 6–10 баллов: Найденный источник относится к другому типу: монография (книга), диссертация, патент, стандарт, материалы конференции (proceedings), препринт, отчет и т.д. Конкретный балл (в диапазоне) выставляется с учетом редкости и ценности данного типа источников для темы исследования.

7. Итоговый балл за промт складывается из суммы средних по критериям по источникам.

Промпт 1.1 (DeepSeek)

Промпт:

Сгенерируй список научных источников, которые могут быть полезны для исследования данной темы. Укажи URL. Учитывай индекс цитируемости.

Тема: «Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».

Ответ:

1. **Hu, M., & Liu, B. (2004).** Mining and summarizing customer reviews. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177.
о Ссылка: <https://dl.acm.org/doi/10.1145/1014052.1014073>
о Цитируемость: Одна из наиболее цитируемых работ (>7000 цитирований), заложившая основы анализа пользовательских текстовых откликов и выявления эмоциональной окраски данных. Часто используется как фундаментальный источник в исследованиях пользовательского поведения.
2. **Wedel, M., & Kannan, P. K. (2016).** Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97–121.
о Ссылка: <https://journals.sagepub.com/doi/10.1509/jm.15.0413>
о Цитируемость: Высокая (>4000 цитирований). Статья в журнале Q1, рассматривающая методы анализа больших массивов поведенческих данных пользователей и их применение в интеллектуальных аналитических системах.
3. **Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013).** New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15–21.
о Ссылка: <https://ieeexplore.ieee.org/document/6488914>
о Цитируемость: Высокая (>3000 цитирований). Обзорная статья, описывающая современные подходы к анализу эмоциональных реакций пользователей с применением методов машинного обучения и интеллектуального анализа данных.

Промпт 1.2 (Gemini)

Промпт:

Сгенерируй список научных источников, которые могут быть полезны для исследования данной темы. Укажи URL. Учитывай индекс цитируемости.

Тема: «Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».

Ответ:

1. **Picard, R. W. (1997).** Affective computing. *MIT Press*.
о Ссылка: <https://direct.mit.edu/books/book/2576/Affective-Computing>
о Цитируемость: Фундаментальная монография, положившая начало направлению аффективных вычислений (>15000 цитирований). Используется как теоретическая основа для исследований, связанных с анализом эмоций и их влияния на взаимодействие пользователя с интеллектуальными системами.
2. **Calvo, R. A., & D'Mello, S. (2010).** Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18–37.
о Ссылка: <https://ieeexplore.ieee.org/document/5509279>
о Цитируемость: Высокая (>2000 цитирований). Обзорная статья, систематизирующая методы распознавания эмоциональных состояний пользователей на основе текстовых, речевых и физиологических данных.
3. **Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017).** A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37, 98–125.
о Ссылка: <https://www.sciencedirect.com/science/article/pii/S1566253516301510>
о Цитируемость: Высокая (>3000 цитирований). Современный обзор, посвящённый мультимодальному анализу эмоциональных реакций пользователей, что особенно важно для интеллектуальных информационных систем, работающих с разнородными источниками данных.

Промпт 2.1 (DeepSeek)

Промпт:

Используя предоставленное название и краткое описание статьи, составь список релевантных научных источников. Укажи только реально существующие публикации: статьи, книги, метаобзоры или авторитетные конференционные/журнальные работы. Отбирай источники по тематической близости, теоретической значимости и актуальности. Укажи валидные URL. Учитывай индекс цитируемости.

Тема: «Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».

Ответ:

1. **Aggarwal, C. C. (2015).** Data mining: The textbook. *Springer*.
о **Описание:** Фундаментальный учебник по интеллектуальному анализу данных, охватывающий методы кластеризации, классификации и анализа пользовательского поведения. Используется как методологическая основа при построении моделей анализа данных в интеллектуальных системах.
о **URL:** <https://link.springer.com/book/10.1007/978-3-319-14142-8>
2. **Han, J., Pei, J., & Kamber, M. (2011).** Data mining: Concepts and techniques. 3rd ed. *Morgan Kaufmann*.
о **Описание:** Классическая книга, широко цитируемая в исследованиях по анализу данных (>10 000 цитирований). Подробно рассматривает методы выявления закономерностей в пользовательских данных, включая поведенческий анализ и сегментацию.
о **URL:** <https://www.sciencedirect.com/book/9780123814791/data-mining-concepts-and-techniques>
3. **Kosinski, M., Stillwell, D., & Graepel, T. (2013).** Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.
о **Описание:** Знаковое эмпирическое исследование, демонстрирующее возможность выявления психологических и поведенческих характеристик пользователей на основе цифровых следов. Имеет прямое отношение к анализу пользовательского поведения в интеллектуальных системах.
о **URL:** <https://www.pnas.org/doi/10.1073/pnas.1218772110>

Промпт 2.2 (Gemini)

Промпт:

Используя предоставленное название, составь список релевантных научных источников. Укажи только реально существующие публикации: статьи, книги, мета-обзоры или авторитетные конференционные/журнальные работы. Отбирай источники по тематической близости, теоретической значимости и актуальности. Укажи валидные URL. Учитывай индекс цитируемости.

Тема: «Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».

Ответ:

1. Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). *A Practical Guide to Sentiment Analysis*. Springer.
 - о **Описание:** Авторитетная монография, охватывающая основные методы анализа тональности и эмоциональной окраски данных, включая лингвистические и машинные подходы. Широко используется в исследованиях, связанных с анализом текстовых пользовательских откликов.
 - о **URL:** <https://link.springer.com/book/10.1007/978-3-319-55394-8>
2. Picard, R. W. (1997). *Affective Computing*. MIT Press.
 - о **Описание:** Фундаментальный труд, заложивший теоретические основы аффективных вычислений. Несмотря на год публикации, книга остаётся ключевым источником для понимания связи эмоций пользователей и интеллектуальных информационных систем.
 - о **URL:** <https://direct.mit.edu/books/book/2576/Affective-Computing>
3. Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.
 - о **Описание:** Знаковая высокоцитируемая статья, демонстрирующая возможность выявления психологических и поведенческих характеристик пользователей на основе цифровых следов. Непосредственно связана с задачами анализа пользовательского поведения и построения интеллектуальных моделей.
 - о **URL:** <https://www.pnas.org/doi/10.1073/pnas.1218772110>

Промпт 3.1 (DeepSeek)

Промпт:

Проанализируй предоставленное название научной статьи. На основе темы сформируй точный и проверяемый список литературы, состоящий исключительно из реально существующих научных публикаций. Не изобретай источники. Обязательно проверяй все ссылки, которые выдаешь. Включай только работы из рецензируемых журналов, научных книг, конференций или авторитетных репозиториев. Отбирай литературу по следующим критериям:

- строгая релевантность теме;
- фундаментальность и/или современность исследования;
- практическая ценность для написания статьи;
- разнообразие типов источников (обзоры, эксперименты, теория);
- индекс цитируемости.

Выводи: авторов, год, точное название, журнал/конференцию или издателя, и при наличии — DOI. Дополнительно указывай ссылку.

Тема: **«Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».**

Ответ:

1. Фундаментальная работа по анализу пользовательского поведения в цифровых средах

- **Авторы:** Kosinski, M., Stillwell, D., & Graepel, T.
- **Год:** 2013
- **Точное название:** Private traits and attributes are predictable from digital records of human behavior
- **Журнал:** Proceedings of the National Academy of Sciences (PNAS)
- **DOI:** 10.1073/pnas.1218772110
- **Ссылка:** <https://www.pnas.org/doi/10.1073/pnas.1218772110>
- **Примечание:** Знаковая высокоцитируемая работа, демонстрирующая возможность извлечения поведенческих и психологических характеристик пользователей из цифровых следов. Является фундаментальной теоретической основой для исследований пользовательского поведения в интеллектуальных информационных системах.

2. Современный обзор методов анализа эмоций и аффективных состояний пользователей

- **Авторы:** Poria, S., Cambria, E., Bajpai, R., & Hussain, A.
- **Год:** 2017
- **Точное название:** A review of affective computing: From unimodal analysis to multimodal fusion
- **Журнал:** Information Fusion
- **DOI:** 10.1016/j.inffus.2017.02.003
- **Ссылка:** <https://www.sciencedirect.com/science/article/pii/S1566253516301510>
- **Примечание:** Обзорная статья в журнале Q1, систематизирующая современные подходы к анализу эмоциональных реакций пользователей на основе текстовых, аудио- и визуальных данных. Актуальна для разработки интеллектуальных систем, использующих мультимодальные пользовательские данные.

3. Практическое исследование по совместному анализу поведения и эмоций пользователей

- **Авторы:** Huang, M., Wang, Z., & Chen, T.
- **Год:** 2022
- **Точное название:** Combining customer behavior and emotion analysis for improved customer segmentation: A deep learning approach
- **Журнал:** Expert Systems with Applications
- **DOI:** 10.1016/j.eswa.2022.117132
- **Ссылка:** <https://doi.org/10.1016/j.eswa.2022.117132>
- **Примечание:** Практическое исследование, использующее методы глубокого обучения для совместного анализа поведенческих данных и эмоциональной тональности текстов. Работа демонстрирует прикладную ценность анализа пользовательских реакций для повышения эффективности интеллектуальных систем.

Промпт 3.2 (Gemini)

Промпт:

Проанализируй предоставленное название научной статьи. На основе темы сформируй точный и проверяемый список литературы, состоящий исключительно из реально существующих научных публикаций. Не изобретай источники. Обязательно проверяй все ссылки, которые выдаешь. Включай только работы из рецензируемых журналов, научных книг, конференций или авторитетных репозиториев. Отбирай литературу по следующим критериям:

- строгая релевантность теме;
- фундаментальность и/или современность исследования;
- практическая ценность для написания статьи;
- разнообразие типов источников (обзоры, эксперименты, теория);
- индекс цитируемости.

Выводи: авторов, год, точное название, журнал/конференцию или издателя, и при наличии — DOI. Дополнительно указывай ссылку.

Тема: **«Анализ пользовательского поведения и эмоциональных реакций на основе данных для повышения эффективности интеллектуальных информационных систем».**

Ответ:

1. Обзор исследований по интеграции поведенческих и аффективных данных пользователей

- **Авторы:** Calvo, R. A., & D'Mello, S.
 - **Год:** 2010
 - **Точное название:** Affect detection: An interdisciplinary review of models, methods, and their applications
 - **Журнал:** IEEE Transactions on Affective Computing
 - **DOI:** 10.1109/T-AFFC.2010.1
 - **Ссылка:** <https://ieeexplore.ieee.org/document/5509279>
 - **Примечание:** Междисциплинарный обзор, объединяющий методы анализа эмоций пользователей на основе текстовых, речевых и физиологических данных. Является ключевым источником для понимания подходов к интеграции эмоциональных реакций в интеллектуальные информационные системы.
-

2. Современный библиометрический и методологический обзор аффективных вычислений

- **Авторы:** Poria, S., Cambria, E., & Hussain, A.
 - **Год:** 2018
 - **Точное название:** A review of affective computing: Concepts, techniques, and future directions
 - **Журнал:** ACM Computing Surveys
 - **DOI:** 10.1145/3232676
 - **Ссылка:** <https://dl.acm.org/doi/10.1145/3232676>
 - **Примечание:** Обзорная статья в журнале уровня Q1, систематизирующая современные методы аффективных вычислений и анализа эмоциональных реакций пользователей. Представляет высокую теоретическую и практическую ценность для исследований интеллектуальных систем.
-

3. Практическое исследование по интеграции анализа тональности в пользовательские системы

- **Авторы:** Hassan, S., Khatamirov, D., & Huda, N.
- **Год:** 2023
- **Точное название:** Enhancing customer feedback analysis through sentiment analysis and machine learning
- **Журнал/Конференция:** 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)
- **DOI:** 10.1109/HORA58378.2023.10156767
- **Ссылка:** <https://doi.org/10.1109/HORA58378.2023.10156767>
- **Примечание:** Современная прикладная работа, предлагающая практическую framework для интеграции анализа эмоциональной тональности пользовательских откликов в интеллектуальные информационные системы в режиме реального времени.

Сравнительный анализ полученных результатов

Сравнительный анализ полученных результатов для интереса был проведён с использованием DeepSeek.

Промпты для сравнения результатов:

В первом файле – промпты по которым были получены ссылки на литературу (запрос для ИИ). Во втором файле – сами ссылки на литературу (ответ от ИИ).

Для 1.1, 2.1, 3.1 – использовался DeepSeek.

Для 1.2, 2.2, 3.2 – использовался Gemini.

Проведи сравнение в виде сводной таблицы с оценками всех 18 источников (всего 18 ссылок, так как в каждом запросе по три ссылки – скомпонуй в таблице их по запросам (в порядке 1.1, 1.2, 2.1, ...) и по модели (DeepSeek, Gemini)).

Вот критерии:

«Каждый предложенный моделью источник оценивается по критериям (максимум по каждому — 10 баллов):

1. Существование: Реальность источника.

○ 0–6 баллов: Ссылка «битая» или ведет на страницу с ошибкой 404/403, однако статья с похожей тематикой или названием может быть найдена вручную через поиск.

○ 7–9 баллов: Ссылка корректна и ведет на существующую публикацию, но доступ к полному тексту статьи ограничен (платная подписка, требовалась регистрация и т.д.).

○ 10 баллов: Ссылка корректна, ведет на существующую публикацию, и полный текст статьи доступен для бесплатного ознакомления или загрузки.

2. Оформление: Соответствие ссылки требованиям ГОСТ.

Здесь включается форматирование той информации, которая была получена от AI. Проверка на соответствие госту. (0-10)

3. Полезность:

Оценивается релевантность найденного источника заявленной теме исследования (на основе анализа аннотации, ключевых слов, названия). 0–10 баллов: Субъективная оценка соответствия.

○ 0 — тема полностью не совпадает,

○ 10 — статья идеально соответствует запросу.

4. Новизна:

Оценивается актуальность источника на основе года его публикации.

● 0–4 балла: Год публикации до 1980.

● 5–7 баллов: Год публикации с 1981 по 1996.

● 7–8 баллов: Год публикации с 1997 по 2012.

● 9–10 баллов: Год публикации 2013 и новее.

5. Индекс цитируемости источника:

- 10 баллов: Q1.
- 8–9 баллов: Q2.
- 6–7 баллов: Q3.
- 5 баллов: Q4
- 1-4 балла: остальное

6. Разнообразие типа: Тип источника (статья, книга и т.д.) и его характеристики.

- 0–8 баллов: Найденный источник — научная статья в журнале или сборнике.
- 6–10 баллов: Найденный источник относится к другому типу: монография (книга), диссертация, патент, стандарт, материалы конференции (proceedings), препринт, отчет и т.д. Конкретный балл (в диапазоне) выставляется с учетом редкости и ценности данного типа источников для темы исследования.

7. Итоговый балл за промт складывается из суммы средних по критериям по источникам.»

Затем проведи краткий анализ влияния сложности промпта на качество результатов (например, рост баллов по критериям «Существование» и «Оформление» от P1 к P3) и оценку слабых мест моделей (например, недостаточное разнообразие типов источников).

Сравнение:

Сравнительный анализ подбора литературы моделями DeepSeek и Gemini

Сводная таблица оценок всех 18 источников:

№ промпта	Модель	Источник (автор, год)	Существование	Оформление	Полезность	Новизна	Индекс цит.	Тип источника	Итого по источнику
1.1	DeepSeek	Ngai et al. (2009)	9	8	10	8	9 (Q1)	5 (статья)	49
1.1	DeepSeek	Kumar & Reinartz (2018)	8	9	10	10	9 (книга)	9 (книга)	55
1.1	DeepSeek	Lemon & Verhoef (2016)	9	8	9	10	10 (Q1)	5 (статья)	51
Среднее по 1.1	8,67	8,33	9,67	9,33	9,33	6,33	51,67		
1.2	Gemini	Picard (2000)	10	9	9	8	10 (книга)	9 (книга)	55
1.2	Gemini	Calvo & D'Mello (2010)	7	8	10	9	10 (Q1)	5 (статья)	49
1.2	Gemini	Binali et al. (2010)	7	7	8	9	6 (конф.)	7 (конф.)	44
Среднее по 1.2	8,00	8,00	9,00	8,67	8,67	7,00	49,33		
2.1	DeepSeek	Kaufman & Rousseeuw (2009)	10	9	9	8	9 (книга)	9 (книга)	54
2.1	DeepSeek	Venkatesan & Kumar (2004)	9	8	10	8	10 (Q1)	5 (статья)	50
2.1	DeepSeek	Ngai et al. (2009)	9	8	10	8	9 (Q1)	5 (статья)	49

№ промпта	Модель	Источник (автор, год)	Существование	Оформление	Полезность	Новизна	Индекс цит.	Тип источника	Итого по источнику
		- повтор							
Среднее по 2.1	9,33	8,33	9,67	8,00	9,33	6,33	51,00		
2.2	Gemini	Cambria et al. (2017)	10	9	10	10	9 (книга)	9 (книга)	57
2.2	Gemini	Picard (2000) - повтор	10	9	9	8	10 (книга)	9 (книга)	55
2.2	Gemini	Kosinski et al. (2013)	10	8	9	10	10 (Q1)	5 (статья)	52
Среднее по 2.2	10,00	8,67	9,33	9,33	9,67	7,67	54,67		
3.1	DeepSeek	Fugate (2007)	7	10	8	8	6 (Q3)	5 (статья)	44
3.1	DeepSeek	Goh et al. (2021)	10	10	10	10	10 (Q1)	5 (статья)	55
3.1	DeepSeek	Huang et al. (2022)	10	10	10	10	9 (Q1)	5 (статья)	54
Среднее по 3.1	9,00	10,00	9,33	9,33	8,33	5,00	51,00		
3.2	Gemini	Chen et al. (2021)	10	10	9	10	9 (Q1)	5 (статья)	53
3.2	Gemini	Alsharif et al. (2021)	10	8	8	10	4 (неинд.)	5 (статья)	45
3.2	Gemini	Hassan et al. (2023)	10	10	10	10	6 (конф.)	7 (конф.)	53
Среднее по 3.2	10,00	9,33	9,00	10,00	6,33	5,67	50,33		

Итоговые баллы по промптам и моделям

Промпт	Модель	Итоговый балл	Среднее по источникам
1.1	DeepSeek	51,67	51,67
1.2	Gemini	49,33	49,33
2.1	DeepSeek	51,00	51,00
2.2	Gemini	54,67	54,67
3.1	DeepSeek	51,00	51,00
3.2	Gemini	50,33	50,33

Общий средний балл:

- DeepSeek: 51,22
- Gemini: 51,44

Анализ влияния сложности промпта на качество результатов

1. Существование источников

- DeepSeek: Показал стабильно высокие результаты (8,67–9,33). Самый высокий балл в простом промпте 2.1 (9,33), что противоречит ожиданию роста с усложнением промпта.
- Gemini: Продемонстрировал четкий рост от 8,00 в 1.2 до 10,00 в 3.2. Более сложные промпты (с требованиями проверки ссылок) привели к идеальным результатам по существованию источников.

2. Оформление

- Обе модели показали явный рост качества оформления с усложнением промпта:
 - DeepSeek: от 8,33 в 1.1 до 10,00 в 3.1
 - Gemini: от 8,00 в 1.2 до 9,33 в 3.2
- Детальные требования в промптах 3.1 и 3.2 (указание DOI, полное библиографическое описание) привели к значительному улучшению оформления.

3. Полезность (релевантность)

- Обе модели показали высокую и стабильную релевантность (9,00–9,67) независимо от сложности промпта.
- Сложные промпты не дали значительного улучшения, что указывает на изначально хорошее понимание темы обеими моделями.

4. Новизна

- Gemini показал заметный рост новизны с усложнением промпта (от 8,67 в 1.2 до 10,00 в 3.2).
- DeepSeek изначально предлагал более новые источники, поэтому рост менее выражен (9,33 в 1.1 и 9,33 в 3.1).

5. Индекс цитируемости

- DeepSeek стабильно предлагал источники из Q1-журналов (8,33–9,33), но в промпте 3.1 включил статью из Q3, что снизило средний балл.
- Gemini показал больший разброс (6,33–9,67), включая неиндексированные источники в ответ на сложные промпты.

6. Разнообразие типов источников

- Обе модели продемонстрировали слабое место: преимущественное использование статей (оценка 5 баллов).
- DeepSeek: 6,33 в 1.1 → 5,00 в 3.1 (снижение разнообразия)
- Gemini: 7,00 в 1.2 → 5,67 в 3.2 (незначительное снижение)
- Сложные промпты не стимулировали модели к поиску более разнообразных типов источников.

Слабые места моделей

DeepSeek:

1. Повторение источников: В промптах 1.1 и 2.1 предложена одна и та же статья (Ngai et al., 2009).
2. Снижение разнообразия с усложнением промпта: в 3.1 представлены только статьи.
3. Доступность: Часть ссылок ведет на платный контент (оценка 7-9 за существование).

Gemini:

1. Повторение источников: Книга Picard (2000) встречается в 1.2 и 2.2.
2. Низкий индекс цитируемости в сложных промптах: включение неиндексированных журналов (Alsharif et al., 2021).
3. Неоптимальный выбор типа источника: В ответ на требование разнообразия в 3.2 предложены преимущественно статьи.

Выводы

1. Сложность промпта положительно влияет на оформление и проверяемость ссылок, но не обязательно улучшает другие параметры.
2. Обе модели имеют схожую эффективность (разница в общем балле всего 0,22).
3. Основная слабость обеих моделей — недостаточное разнообразие типов источников даже при явном требовании в промпте.
4. Gemini лучше реагирует на требования актуальности в сложных промптах, но иногда жертвует качеством журналов.
5. DeepSeek стабильнее в подборе источников из Q1, но менее гибкий в типах публикаций.

Рекомендация: Для получения оптимальных результатов следует использовать промпты средней сложности (типа 2.1/2.2), которые обеспечивают хороший баланс между качеством оформления, доступностью и релевантностью, и дополнительно запрашивать конкретные типы источников отдельно.