

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

Выполнил
студент гр. № 5040102/50201

Худина А.А.

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:

Сгенерируй список научных источников, которые могут быть полезны для исследования данной темы. Укажи URL. Учитывай индекс цитируемости. Тема: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. **Ngai, E.W.T., Xiu, L., & Chau, D.C.K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36(2), 2592-2602.**
 - **Ссылка:** <https://www.sciencedirect.com/science/article/abs/pii/S0957417408001223>
 - **Цитируемость:** Очень высоко цитируемая обзорная статья (>2000 цитирований), классика в области data mining для CRM. Показывает методы и направления.
2. **Kumar, V., & Reinartz, W. (2018). Customer Relationship Management: Concept, Strategy, and Tools. 3rd ed. Springer.**
 - **Ссылка:** <https://link.springer.com/book/10.1007/978-3-662-55381-7>
 - **Цитируемость:** Фундаментальный современный учебник. Глава о сегментации на основе клиентского жизненного цикла и ценности критически важна. Книга широко цитируется в академических работах.
3. **Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96.**
 - **Ссылка:** <https://journals.sagepub.com/doi/abs/10.1509/jm.15.0420>
 - **Цитируемость:** Высокая (>5000 цитирований). Статья в топ-журнале (Q1) задает рамки анализа поведения через призму customer journey, что является ключом к современной CRM.

Промпт:

Сгенерируй список научных источников, которые могут быть полезны для исследования данной темы. Укажи URL. Учитывай индекс цитируемости. Тема: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. **Picard, R. W. (2000). Affective computing.** MIT press.
 - Ссылка: <https://direct.mit.edu/books/book/2579/Affective-Computing>
 - Цитируемость: Краеугольный труд, основавший направление аффективных вычислений (>15000 цитирований). Обязателен для теоретического обоснования.
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.
 - Ссылка: <https://ieeexplore.ieee.org/document/5509279>
 - Цитируемость: Высокая (>2000 цитирований). Систематический обзор методов детекции аффекта (текст, речь, физиология, лицо) в ведущем журнале по теме.
3. **Binali, H., Wu, C., & Potdar, V. (2010). A new significant area: Emotion detection in E-Learning using opinion mining techniques.** *In 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*.
 - Ссылка: <https://ieeexplore.ieee.org/document/5565060>
 - Цитируемость: Хотя это конференция, работа часто цитируется как пример применения анализа тональности (opinion mining/sentiment analysis) для выявления эмоций в текстовых данных (отзывы, чаты, соцсети) — самом доступном для CRM источнике.

Промпт:

Используя предоставленное название и краткое описание статьи, 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: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. **Kaufman, L., & Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons.**
 - **Описание:** Классическая книга по кластерному анализу, рассматривающая основные алгоритмы (k-means, hierarchical, DBSCAN и др.), их выбор и оценку качества кластеризации. Необходима для методологического фундамента.
 - **URL:** <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316801> (DOI)
2. **Венкатесан, Р., и Кумар, В. (2004). Модель ценности клиента: подход к управлению взаимоотношениями с клиентами, основанный на данных. *Journal of Marketing*, 68(1), 106-125.**
 - **Описание:** Одна из самых цитируемых статей в области CRM, вводит концепцию Customer Lifetime Value (CLV) и предлагает основу для сегментации на основе ценности, а не только демографии.
 - **URL:** <https://journals.sagepub.com/doi/10.1509/jmkg.68.1.106.24028>
3. **Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36(2), 2592-2602.**
 - **Описание:** Обзорная статья, систематизирующая применение методов data mining (включая кластеризацию) в CRM. Хорошая отправная точка для понимания ландшафта исследований.
 - **URL:** <https://www.sciencedirect.com/science/article/abs/pii/S0957417408002874>

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: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. **Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). *A Practical Guide to Sentiment Analysis*. Springer.**
 - **Описание:** Авторитетное руководство по анализу тональности, включая лингвистические и машинные подходы. Ключевой источник для анализа текстовых эмоциональных откликов (отзывы, поддержка, соцсети).
 - **URL:** <https://link.springer.com/book/10.1007/978-3-319-55394-8>
2. **Picard, R. W. (2000). *Affective Computing*. MIT Press.**
 - **Описание:** Фундаментальная книга, вводящая понятие "аффективных вычислений". Хотя часть технологий устарела, теоретическая база для понимания связи эмоций и данных остается актуальной.
 - **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.**
 - **Описание:** Знаковое исследование, демонстрирующее, как цифровое поведение (лайки в Facebook) может предсказывать психологические черты. Прямо связано с идеей сегментации на основе глубинных характеристик.
 - **URL:** <https://www.pnas.org/doi/10.1073/pnas.1218772110>

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. Если тема широкая, включи ключевые базовые источники и последние значимые работы. Старайся охватить основные направления, необходимые для полноценного академического исследования. Ко всем статьям, которые ты предложишь, укажи валидные ссылки, по которым можно прочитать полные версии текстов. Тема: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. Фундаментальная работа по связи эмоций, поведения и лояльности (базовый источник)

- **Авторы:** Fugate, D. L.
- **Год:** 2007
- **Точное название:** Neuromarketing: a layman's look at neuroscience and its potential application to marketing practice.
- **Журнал/Издатель:** Journal of Consumer Marketing
- **DOI:** 10.1108/07363760710834808
- **Ссылка:** <https://doi.org/10.1108/07363760710834808>
- **Примечание:** Хотя статья не новая, она является часто цитируемой фундаментальной работой, обосновывающей важность нейронауки и эмоциональных откликов для понимания потребительского поведения, что служит теоретической базой для современных подходов.

2. Современный обзор методов анализа эмоций в потребительских данных

- **Авторы:** Goh, K.-Y., Chu, H.-C., & Wu, J.-J.
- **Год:** 2021
- **Точное название:** Affective analytics of customers: A review and research agenda.
- **Журнал/Издатель:** International Journal of Research in Marketing
- **DOI:** 10.1016/j.ijresmar.2020.12.002
- **Ссылка:** <https://doi.org/10.1016/j.ijresmar.2020.12.002>

- **Примечание:** Актуальный обзорный article в топовом маркетинговом журнале (Q1). Анализирует современные методы (текстовый, аудио, видео, физиологический анализ) для измерения аффективных состояний клиентов, что напрямую относится к «анализу эмоциональных откликов».

3. Практическое исследование по сегментации на основе цифрового поведения и эмоций в тексте

- **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/Конференция:** Expert Systems with Applications
- **DOI:** 10.1016/j.eswa.2022.117132
- **Ссылка:** <https://doi.org/10.1016/j.eswa.2022.117132>
- **Примечание:** Практическое исследование, применяющее методы глубокого обучения (LSTM, Attention) для совместного анализа поведения на сайте и эмоциональной тональности отзывов/сообщений. Прямая иллюстрация «подхода к сегментации».

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. Если тема широкая, включи ключевые базовые источники и последние значимые работы. Старайся охватить основные направления, необходимые для полноценного академического исследования. Ко всем статьям, которые ты предложишь, укажи валидные ссылки, по которым можно прочитать полные версии текстов. Тема: «Разработка подхода к сегментации пользователей на основе анализа их поведения и эмоциональных откликов для повышения эффективности CRM-систем».

Ответ:

1. Исследование по интеграции психографических и поведенческих данных в CRM

- **Авторы:** Chen, Y., Mandler, T., & Meyer-Waarden, L.
- **Год:** 2021
- **Точное название:** Three decades of research on loyalty programs: A literature review and future research agenda.
- **Журнал/Издатель:** Journal of Business Research
- **DOI:** 10.1016/j.jbusres.2021.03.038
- **Ссылка:** <https://doi.org/10.1016/j.jbusres.2021.03.038>
- **Примечание:** Ключевой обзор, связывающий различные типы данных о клиентах (включая поведенческие и, косвенно, аффективные) с мерами повышения лояльности в рамках CRM-стратегий. Объясняет «повышение эффективности CRM-систем».

2. Экспериментальное исследование по использованию физиологических данных (нейромаркетинг) для сегментации

- **Авторы:** Alsharif, A. H., Salleh, N. Z. M., & Baharun, R.
- **Год:** 2021
- **Точное название:** A bibliometric analysis of neuromarketing: Current status and future directions.
- **Журнал/Издатель:** Journal of Theoretical and Applied Information Technology
- **Ссылка:** <http://www.jatit.org/volumes/Vol99No19/10Vol99No19.pdf>

- **Примечание:** Хотя это обзор, он предоставляет исчерпывающую и современную картину методов нейромаркетинга (EEG, GSR, eye-tracking) для измерения эмоциональных и когнитивных откликов. Указан как ключевой источник для понимания инструментария анализа эмоций. Полный текст доступен по прямой ссылке.

3. Работа, посвященная операционализации эмоционального профиля клиента в CRM

- **Авторы:** Hassan, S., Khantimirov, D., & Huda, N.
- **Год:** 2023
- **Точное название:** Enhancing Customer Relationship Management through Sentiment Analysis and Machine Learning: A Framework for Real-Time Customer Feedback Integration.
- **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 работа на конференции, предлагающая конкретную framework для интеграции анализа эмоциональной тональности (sentiment analysis) в реальном времени в CRM-системы. Отвечает критерию практической ценности и современности.

Сравнительный анализ полученных результатов

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

Промпт 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: Проявил 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.