

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ

Тема: Генерация списка литературы

Выполнила: Буталова Юлия

Замечания по ЛР: немного нецелесообразно выбирать 18 различных источников. Мне кажется, было бы более удачно сравнить по 9 источников из каждой модели, чтобы выяснить, к какой модели больше доверия. То есть, в моих запросах у двух моделей были некоторые одинаковые источники, но из-за того, что нужно собрать 18 «универсальных источников» я не могла взять хорошие источники от второй модели, так как уже взяла их от первой, что портит итоговую статистику. Так же в задании сначала написано, что нужно выбрать «живые» источники, а потом в одном из критериев написано, что нужно указать, живой источник или нет – конечно живой, я же в самом начале выбирала именно живые, в чем смысл этого критерия, не пойму.

Промпты:

P1

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

P2

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.

P3

Проанализируй предоставленные название и краткое описание 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

Выводы: авторов, год, точное название, журнал/конференцию или издателя, и при наличии — DOI. Если тема широкая, включи ключевые базовые источники и последние значимые работы. Страйся 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: КЛАСТЕРИЗАЦИЯ РУКОПИСНЫХ ЦИФР НЕЙРОННОЙ СЕТЬЮ
КОХОНЕНА**

Описание: Проведена кластеризация рукописных цифр по шестидесяти тысячам изображений, содержащихся в обучающей выборке базы данных MNIST. Для кластеризации использована нейронная сеть Кохонена. Для каждой рукописной цифры определено оптимальное количество кластеров (не более 50). При определении расстояния между объектами (изображениями рукописных цифр) использована евклидова норма. Проверка правильности построения кластеров проведена по данным из тестовой выборки базы MNIST. Тестовая выборка содержит десять тысяч изображений. Сделан вывод о том, что изображения из тестовой выборки принадлежат кластеру «правильной цифры» с вероятностью более 90%. Для каждой цифры вычислена F-мера для оценки кластеров. Наилучшие значения F-меры получены для цифр 0 и 1 (F-среднее значение равно 0,974). Наихудшие значения получены для цифры 9 (F-среднее значение равно 0,903). Также проведен кластерный анализ, который позволил сделать выводы о возможных ошибках в распознавании нейронной сетью Кохонена. Построены пересечения кластеров для изображений рукописных цифр. Приведены примеры пересечений кластеров, а также показаны примеры изображений, которые распознаны нейронной сетью неверно.

GEMINI 3 PRO

Результат первого промпта

1. Kussul, E. M., & Baidyk, T. N. (2004). Improved method of handwritten digit recognition tested on MNIST database. *Image and Vision Computing*, 22(12), 971–981.

Описание: Работа, в которой рассматриваются методы улучшения распознавания на MNIST, что полезно для сравнения с вашими результатами (особенно касательно ошибок на цифре 9).

URL: <https://doi.org/10.1016/j.imavis.2004.03.008>

2. Deselaers, T., Keysers, D., & Ney, H. (2008). Features for image retrieval: an experimental comparison. *Information Retrieval*, 11(2), 77–107.

Описание: Хотя статья про поиск изображений, она содержит глубокий анализ метрик расстояния (включая Евклидову) применительно к базе MNIST и оценку качества кластеризации (F-measure, Precision/Recall).

URL: <https://link.springer.com/article/10.1007/s10791-007-9039-3>

3. Xie, J., Girshick, R., & Farhadi, A. (2016). Unsupervised Deep Embedding for Clustering Analysis. *International Conference on Machine Learning (ICML)*.

Описание: Одна из ключевых современных работ по кластеризации MNIST (алгоритм DEC). Полезна для раздела "Обзор литературы" или "Сравнение

результатов", чтобы показать место SOM среди современных методов глубокого обучения.

URL: <http://proceedings.mlr.press/v48/xieb16.pdf>

Результат второго промпта

4. Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37, 52–65. * **Почему это важно:** Более поздняя работа автора метода, в которой он подводит итоги и уточняет математические аспекты, включая использование евклидовой нормы для определения расстояния между объектами, что прямо указано в вашем описании. * **URL:** <https://doi.org/10.1016/j.neunet.2012.09.018>

5. Pöhlbauer, G., Dittenbach, M., & Rauber, A. (2006). Advanced visualization of Self-Organizing Maps with vector fields. *Neural Networks*, 19(6-7), 911–922. * **Почему это важно:** Статья посвящена визуализации границ кластеров и анализу их пересечений. Это крайне полезно для вашей части работы, где вы описываете «пересечения кластеров» и визуальный анализ ошибок. * **URL:** <https://doi.org/10.1016/j.neunet.2006.05.013>

6. Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of Intelligent Information Systems*, 17(2-3), 107–145. * **Почему это важно:** Фундаментальный обзор метрик валидации кластеров. Поскольку вы используете F-меру для оценки качества кластеров и анализа «правильной цифры», эта работа необходима для методологического обоснования выбранного способа оценки.
* **URL:** <https://link.springer.com/article/10.1023/A:1012801612483>

Результат третьего промпта

7. Автор/Год: Kohonen, T. (1990). **Название:** *The self-organizing map*. **Издание:** Proceedings of the IEEE, 78(9), 1464–1480. **DOI:** [10.1109/5.58325](https://doi.org/10.1109/5.58325) **Ссылка на текст:** [IEEE Xplore PDF](#) (или [препринт](#))

Почему включено: Это классическая работа создателя метода Тево Кохонена. Она обязательна для раздела "Методология", где вы описываете алгоритм обучения сети, понятие "нейрона-победителя" и адаптацию весовых коэффициентов. Здесь также обосновывается использование евклидовой метрики для входных векторов, что соответствует вашему описанию.

8. Автор/Год: Vesanto, J., & Alhoniemi, E. (2000). **Название:** *Clustering of the self-organizing map*. **Издание:** IEEE Transactions on Neural Networks, 11(3), 586–600. **DOI:** [10.1109/72.846731](https://doi.org/10.1109/72.846731) **Ссылка на текст:** [ResearchGate Full Text](#)

Почему включено: Ключевая работа, описывающая двухуровневый подход: сначала обучение SOM, затем кластеризация прототипов (весов нейронов). Это напрямую коррелирует с вашей задачей "определения оптимального количества кластеров". Статья также описывает методы визуализации расстояний (U-matrix), что полезно для анализа "пересечения кластеров".

9. Автор/Год: LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). **Название:** *Gradient-based learning applied to document recognition*. **Издание:** Proceedings of the IEEE, 86(11), 2278–2324. **DOI:** [10.1109/5.726791](https://doi.org/10.1109/5.726791) **Ссылка на текст:** [Официальный PDF \(Yann LeCun\)](#)

Почему включено: Первостоиник базы данных MNIST. Статья необходима для описания характеристик выборки (60 000 train / 10 000 test). Авторы приводят детальный анализ морфологии цифр, объясняя, почему цифры "4" и "9" часто путаются (из-за схожего написания верхней петли), что подтверждает ваши выводы о низкой F-мере для цифры 9 (0.903).

GPT 5.1

Результат первого промпта

10. Bishop, C. M.

Pattern Recognition and Machine Learning. Springer, 2006.

(см. главы по кластеризации, SOM, F-мере)

URL: <https://link.springer.com/book/10.1007/978-0-387-45528-0>

11. Jain, A. K., Murty, M. N., Flynn, P. J.

Data clustering: a review. ACM Computing Surveys, 1999.

DOI: 10.1145/331499.331504

URL: <https://dl.acm.org/doi/10.1145/331499.331504>

12. Xu, R., Wunsch, D.

Survey of clustering algorithms. IEEE Transactions on Neural Networks, 2005.

DOI: 10.1109/TNN.2005.845141

URL: <https://ieeexplore.ieee.org/document/1427769>

Результат второго промпта

13. Hastie, T., Tibshirani, R., Friedman, J.

The Elements of Statistical Learning. 2nd ed. Springer, 2009.

(гл. 13–14: кластеризация, оценка качества кластеров)

DOI: 10.1007/978-0-387-84858-7

URL: <https://link.springer.com/book/10.1007/978-0-387-84858-7>

14. Manning, C. D., Raghavan, P., Schütze, H.

Introduction to Information Retrieval. Cambridge University Press, 2008.

(формальное введение precision/recall/F1)

Онлайн-версия: <https://nlp.stanford.edu/IR-book/>

15. Haykin, S.

Neural Networks and Learning Machines. 3rd ed. Pearson, 2008.

(разделы по самоорганизующимся картам, евклидовой норме и векторному квантованию)

Издатель: <https://www.pearson.com/en-us/subject-catalog/p/neural-networks-and-learning-machines/P200000003797>

Результат третьего промпта

16. Jain, A. K., Duin, R. P. W., Mao, J. (2000). Statistical pattern recognition: A review.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1), 4–37.

DOI: 10.1109/34.824819

Полный текст (PDF через IEEE Xplore / библиотеки):

<https://ieeexplore.ieee.org/document/824819>

17. Yan Yang (2011). Asymptotic Convergence Properties of the EM Algorithm for Mixture of Experts, 22(12), 3207–3220.

DOI: 10.1162/NECO_a_00052

Полный текст (PDF у издателя / в институциональных репозиториях):

<https://direct.mit.edu/neco/article/22/12/3207/7691>

18. Cottrell, M., Hammer, B., Hasenfuß, A., Villmann, T. (2006). Batch and incremental learning for vector quantization and self-organizing maps. In: *Knowledge Discovery in Databases: PKDD*, LNCS 4213, 193–205. Springer.

DOI: 10.1007/11871637_20

Полный текст (PDF у Springer):

https://link.springer.com/chapter/10.1007/11871637_20

Оценка источников

Критерии оценивания

- 1.** Существование
- 2.** Оформление
- 3.** Полезность
- 4.** Новизна
- 5.** Индекс цитируемости источника
- 6.** Разнообразие типа

Источник	Крит 1	Крит 2	Крит 3	Крит 4	Крит 5	Крит 6	Средний балл
GEMINI 3 PRO							
1(1)	9	7	8	8	9	8	8,2
2(1)	9	7	7	8	8	8	8,0
3(1)	10	7	8	10	10	9	8,3
							24,5
4(2)	9	7	10	9	10	8	8,7
5(2)	9	7	8	8	10	8	8,0
6(2)	9	7	9	8	9	8	8,5
							25,2
7(3)	9	7	10	7	10	8	8,7
8(3)	9	7	10	8	10	8	8,7
9(3)	10	7	8	8	10	8	8,3
							25,7
GPT 5.1							
10(1)	9	7	8	8	10	9	8,8
11(1)	8	7	8	8	10	8	8,7
12(1)	8	7	9	8	10	8	8,7
							26,2
13(2)	10	7	8	8	10	9	8,7

14(2)	10	7	7	8	10	9	8,5
15(2)	8	7	8	8	9	10	8,7
							25,9
16(3)	8	7	7	8	10	8	8,7
17(3)	9	6	6	8	9	8	8,7
18(3)	9	7	9	8	8	9	8,5
							25,9

1. У модели **GEMINI 3 PRO** усложнение промптов практически не повлияло на Существование и Оформление, но с каждым разом повышало Полезность, Индекс цитируемости, при этом такие показатели как Новизна и Разнообразие типа понижались.

Общая динамика положительная, для данной модели суммарная оценка по всем источникам при увеличении сложности промпта возрасла.

2. У модели **GPT 5.1** частичное улучшение оценки произошло лишь по критериям Существование и Разнообразие типа, по 2 и 4 критериям изменений не произошло, а остальные критерии дали более низкие оценки после усложнения промптов.

Общая динамика отрицательная, усложнение промптов привело к ухудшению суммарных показателей.

Обе модели показали свои слабые места в составлении существующих (настоящих/не удаленных) источников, в следствие того, что обладают данным лишь за 2024 год. Больше половины из выданных на каждом промпте источников либо вели на пустые страницы, либо на страницы с ошибкой. Так же очень распространенной проблемой обеих моделей являлось то, что ссылка вела не на ту статью или книгу, которая указана в источнике.

Выводы:

- Для модели **GEMINI 3 PRO** успешным будет наибольшая детализация, уточнение промпта, внесение дополнительных критериев и параметров, а для модели **GPT 5.1** уточнения только мешают, более лаконичный и чёткий промпт приносит лучшие результаты. Модель хорошо обучена, и сама понимает, как выбирать источники, по каким критериям.
- Слабые места обеих моделей: обе модели отстали от времени, источники, которые они считают актуальными и существующими прекратили своё существование, либо не существовали вовсе.

Замечание про DeepSeek любой версии: изначально планировалась использоваться модель DeepSeek-V3.2. При любом запросе с просьбой найти источники модель отвечала, что не имеет доступа к интернету и не может составить источники, как и аналогичные модели более ранних версий. Так что не представляю, как авторы ЛР6 предполагали использовать её для составления источников