

Сократический метод как инструмент выбора моделей машинного обучения для корпоративных информационных систем

А.А. Касымов¹, А. Лысенко²

¹Воронежский государственный технический университет

²Белгородский государственный аграрный университет имени В.Я. Горина

Аннотация: В статье представлен анализ применения сократического метода для выбора моделей машинного обучения в корпоративных информационных системах. Цель исследования заключается в изучении возможностей использования модульной архитектуры сократической модели для интеграции предобученных моделей без необходимости их дополнительного обучения. Методология основывается на языковом взаимодействии между модулями, что позволяет объединять данные из различных доменов, включая текст, изображения и аудио, для решения многомодальных задач. В результате проведенного ясно, что предложенный подход обладает высоким потенциалом для оптимизации выбора моделей, ускорения процессов принятия решений и снижения затрат на внедрение искусственного интеллекта в корпоративной среде.

Ключевые слова: сократический метод, машинное обучение, корпоративные информационные системы, многомодальные данные, языковое взаимодействие, оптимизация бизнес-процессов, искусственный интеллект.

Введение

В условиях нового времени меняется и корпоративный мир. Искусственный интеллект (ИИ) и машинное обучение (МО) становятся важными инструментами для повышения эффективности работы информационных систем. Ключевым моментом в выборе таких технологий служит тщательный выбор моделей, которые способны решать задачи обработки данных, автоматизации процессов и поддержки принятия решений. На данный момент существует большое количество таких моделей, и все они отличаются своей специализацией. Это, в свою очередь, усложняет процесс выбора конкретной модели, что делает актуальной задачу создания универсальных подходов, позволяющих рационально интегрировать предобученные модели в корпоративные системы [1].

Одним из перспективных методов является сократический подход, представленный в работе [2] под авторством Andy Zenga. В данной концепции взаимодействие между моделями осуществляется с помощью

языковых запросов, что, в свою очередь, позволяет объединять их возможности для выполнения сложных многомодальных задач. Преимущество данного метода значительно, так как он исключает необходимость дополнительного обучения моделей, что снижает затраты и ускоряет их внедрение в реальную практику [1].

Целью данной работы является исследование применения сократического метода для выбора моделей машинного обучения в корпоративных информационных системах. В работе подробно анализируются особенности выбранного метода, его применимость для решения бизнес-задач, а также перспективы оптимизации процессов выбора и внедрения ИИ-технологий.

Материалы и методы

Исследование базируется на методологии Socratic Models [2], метод построен на взаимодействии между моделями через языковые запросы. Основной особенностью подхода является модульная структура, которая позволяет использовать предобученные модели из различных доменов для совместного выполнения задач, при этом каждая модель выполняет свои задачи, и в конце преобразует в общий результат. Этот метод интересен и на первый взгляд очевиден, так как в любом деле каждый имеет свои обязанности, и выполняет то дело, на которое есть квалификация. Так и здесь, так как одна модель нацелена на одну специализацию, другая – на иную. В рамках работы были задействованы модели, такие, как GPT-3 для анализа текстовых данных, CLIP для обработки визуальной информации и BERT для глубокой интерпретации текста. Эти системы взаимодействовали друг с другом, передавая результаты в текстовой форме, что исключает необходимость дополнительного обучения [2].

Для организации работы использовались специально разработанные текстовые шаблоны, которые формируют запросы и упрощают обмен

данными между выбранными моделями. Тестирование производилось на конкретных задачах, связанных с анализом клиентских данных, автоматизацией бизнес-процессов и прогнозированием ключевых показателей. Именно постановка верно поставленных задач позволила увидеть его эффективность в реальных условиях. Взаимодействие моделей в рамках предложенной архитектуры оказалось универсальным, гибким и адаптируемым для решения задач из различных доменов [2].

В исследовании было важно точно выбрать модели, и применить их к разным корпоративным сценариям. Полученные данные подтвердили, что использование языковых запросов для интеграции моделей существенно снижает временные и финансовые затраты на их внедрение, при этом обеспечивая высокую производительность системы.

Литературный обзор

Процесс выбора моделей машинного обучения требует учета множества факторов, включая их специализацию, способность к интеграции и уровень адаптивности. Одной из основополагающих работ в области использования предобученных моделей является исследование [2], которое представляет методологию сократической модели. Их подход основывается на модульной архитектуре, где модели взаимодействуют посредством языковых запросов. Такой метод демонстрирует высокую эффективность при решении задач, связанных с многомодальными данными, объединяя визуальные, текстовые и аудиальные домены.

Ещё важно, что значительный вклад в развитие методов выбора моделей внесли работы Devlin M. [3] и Brown T. [1]. В частности, данные авторы предложили модель BERT, которая стала основой для многих языковых моделей, используемых в корпоративных системах. Их работа указывает на важность предварительного обучения на больших объемах данных и последующего дообучения на специфических задачах. Аналогично,

в [1] разработана модель GPT-3, благодаря ее возможностям в обработке естественного языка, она используется в качестве основы для сократической модели.

Ещё важным направлением, упомянутым в литературе, является объединение различных доменов данных, предложенное Radford A. и другими (2021) [4]. Их работа над моделью CLIP продемонстрировала, как визуальная и текстовая информация могут быть интегрированы для повышения точности и скорости обработки запросов. В рамках исследования Zenga, эта концепция расширена за счет модульной структуры, где такие модели могут эффективно взаимодействовать друг с другом.

Ещё важно отметить исследования Hu and Singh и Jia et al. (2021) [5]. Они предлагают новые алгоритмы, которые способны объединять данные различных типов для достижения более качественных результатов в задачах анализа и прогнозирования. Эти исследования составляют теоретическую базу для разработки и внедрения сократической модели, подчеркивая их универсальность и адаптивность к условиям.

Таким образом, в ходе литературного обзора выявлено, что современная методология выбора моделей машинного обучения основана на использовании предобученных систем и их взаимодействии через языковые запросы. Это подтверждается работами Zenga, Devlin, Brown и других, их труды заложили основу для применения сократического метода в корпоративных информационных системах [1, 2, 6].

Результаты

Применение сократического метода, основанного на методологии Zenga, показало высокую эффективность при решении задач интеграции моделей машинного обучения в корпоративных информационных системах. Исходя из исследования ясно, что использование языкового взаимодействия

между предобученными моделями обеспечивает их синергетический эффект. Несмотря на это, необходимо ещё дальше исследовать данное направление, учитывая потенциальное развитие подхода в направлениях [2].

Визуализация процессов взаимодействия моделей особенно важна для понимания их взаимодополняемости. Пример использования языка для построения истории событий на основе видео можно увидеть на рис.1.

```
Places: staircase. Objects: stairs, animal, mammal, hamster, human leg. Activities: climbing.  
5 Possible Sounds: footsteps, creaking stairs, someone calling your name, a dog barking, a centipede crawling.
```



Рис. 1. – Использование языка для построения истории событий на основе видео [2]

Этот процесс демонстрирует, как сократический метод может быть полезен и внедрен для обработки временных последовательностей, таких, как видеоматериалы, что открывает возможности для более сложных систем управления данными. Автор статьи дает понять, что такая адаптация может быть полезна в бизнес-сценариях, например, для анализа поведенческих

данных пользователей или оптимизации логистики, для построения стратегий, прогнозов и многих других полезных вещей [7].

Ещё немало важно понять, как сократическая модель может конкурировать с другими подходами. По данным рис.2, где приведены метрики эффективности (BLEU-4, METEOR и CIDEr), наглядно видно, что сократическая модель успешно конкурирует с другими подходами. Автор делает акцент на том, что использование таких моделей в условиях ограниченных данных, характерных для многих корпоративных задач, может стать ключевым преимуществом. Даже более того, так как возможность интеграции визуальной интерпретации (например, через CLIP) и текстового анализа (через GPT-3) создает универсальное решение для задач, которые ранее требовали отдельной обработки. За счет этого сокращается время на поиски решения задач, и как итог возникает уникальное решение [4].

Method	BLEU-4	METEOR	CIDEr	SPICE	ROUGE-L
* ClipCap [45]	40.7	30.4	152.4	25.2	60.9
† MAGIC [61]	11.4	16.4	56.2	11.3	39.0
ZeroCap [62]	0.0	8.8	18.0	5.6	18.3
SMs 0-shot (ours)	6.9	15.0	44.5	10.1	34.1
SMs 3-shot (ours)	18.3	18.8	76.3	14.8	43.7

* finetuned on full training set with image-text pairs.

† finetuned on unpaired training set, zero-shot on image-text pairs.

Рис. 2. – Метрики эффективности BLEU-4, METEOR и CIDEr [4]

Стоит взять во внимание также пример взаимодействия моделей для обработки видеоконтента, представленный на рис.3. Здесь четко видна архитектура системы, где каждая модель выполняет свою специализированную функцию, чем говорилось ранее, а итоговый результат формируется на основе текстовых запросов. Это позволяет внедрять

сократическую модель в задачи, связанные с мультимодальными данными, такими как анализ видеонаблюдения или разработка ассистирующих систем в реальном времени [8].

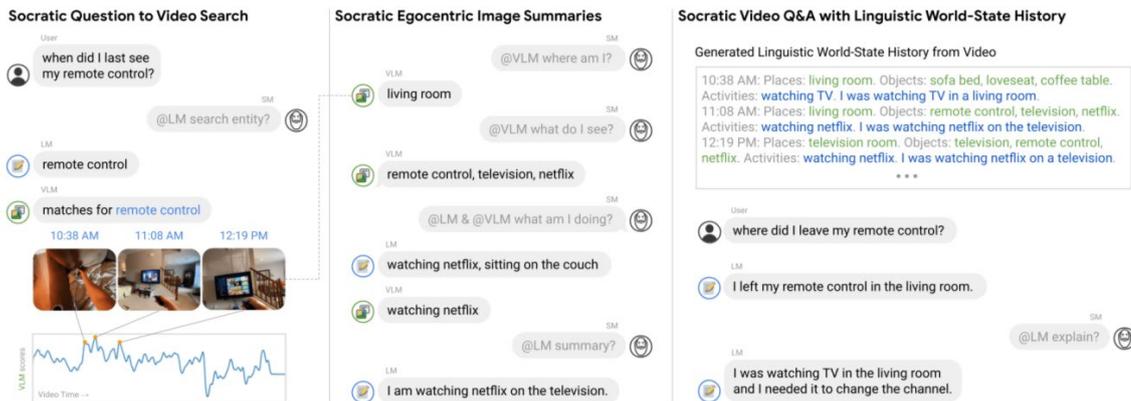


Рис. 3. – Пример взаимодействия моделей для обработки видеоконтента [8]

Авторы работы убеждены, что предложенный подход имеет значительный потенциал для совершенствования, так как алгоритмы и ключевые показатели уже имеются, что дает возможность и дальше развивать данную модель. Например, использование дополнительных слоев оптимизации, таких как контекстные подсказки или кросс-доменное обучение, может повысить точность и адаптивность системы. Это направление, на взгляд автора, остается открытым для дальнейших исследований и интеграции с новыми технологиями, так как с каждым днем появляется все больше новых технологий [9].

Из результатов исследования складывается вывод, который не только подтверждает эффективность сократического метода, но и подчеркивают его перспективность для будущих разработок. Уникальность метода заключается в его модульной структуре, которая позволяет гибко адаптировать предобученные модели к специфическим задачам. Автор уверен, что

дальнейшее развитие данного подхода может стать основой для нового поколения интеллектуальных систем, с чем невозможно не согласиться [10].

Обсуждение

По результатам исследования нельзя не отметить, насколько эффективен сократический метод для интеграции предобученных моделей машинного обучения в корпоративных задачах. Основным преимуществом подхода является возможность объединения моделей различных доменов без необходимости их дообучения, что делает метод особенно актуальным в условиях ограниченных ресурсов и быстро меняющихся условий [11]. Остается только до конца понять, как же правильно для модели формулировать качественные языковые запросы, так как это является основополагающим звеном в данной модели, и ещё важно, что специфика задач, для которой будет использована модель – абсолютно разная, это означает, что языковые запросы для каждой области уникальны и требуют тщательного изучения.

Исследования в данной области будут продолжаться, и дальнейшее совершенствование метода, включая внедрение цепочек подсказок и использование мультимодальных моделей нового поколения, позволит устранить существующие ограничения и расширить сферу применения подхода. Несмотря на существующие ограничения, в корпоративной практике сократический метод уже сегодня способен снизить издержки и ускорить внедрение ИИ, но его реализация требует адаптации процессов управления данными [12].

Заключение

Сократический метод, предложенный в работе Zenga, Maria Attarian, Brian Ichter и других, продемонстрировал свою высокую эффективность как

инструмент выбора и интеграции моделей машинного обучения в корпоративных информационных системах. Это стало возможным благодаря объединению предобученных моделей различных доменов через языковые запросы, что снижает необходимость в дополнительных ресурсах и ускоряет процессы внедрения [6].

В ходе исследования было показано, что предложенный подход способен решать важные задачи в сфере бизнеса, при этом почва для дальнейшего изучения остается объемной. Ключевым вызовом остается настройка системы под конкретные задачи, особенно в условиях быстро меняющихся бизнес-потребностей. Сократический метод является перспективным, он способен решать важные задачи в корпоративном мире, интегрируя искусственный интеллект в него [5].

Таким образом, методология сократического метода предоставляет надежное и гибкое решение для современных вызовов, связанных с выбором и применением моделей машинного обучения. Она открывает возможности для значительного повышения эффективности и конкурентоспособности корпоративных информационных систем уже сейчас, решая важные задачи, но и также языковые запросы остаются актуальной проблемой для дальнейших исследований.

Литература (References)

1. Brown T., Mann B., Ryder N., Subbiah M., Kaplan J. D., Dhariwal P., Amodei D. Language models are few-shot learners // Advances in Neural Information Processing Systems. –2020. V. 33. pp. 187–190.

2. Zeng A., Attarian M., Ichter B., Choromanski K., Wong A., Welker S., Florence P. Socratic Models: Composing zero-shot multimodal reasoning with language. 2022 URL: arXiv preprint arXiv:2204.00598.

3. Devlin J., Chang M.W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. 2018 URL: arXiv preprint arXiv:1810.04805.

4. Radford A., Kim J.W., Hallacy C., Ramesh A., Goh G., Agarwal S., Sutskever I. Learning transferable visual models from natural language supervision // Proceedings of the 38th International Conference on Machine Learning. 2021. pp. 82–87.

5. Xu J., Mei T., Yao T., Rui Y. MSR-VTT: A large video description dataset for bridging video and language // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. pp. 50–52.

6. Devlin J., Chang M.W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding // Proceedings of NAACL-HLT 2019. 2019. pp. 141–143.

7. Yang Z., Gan Z., Wang J., Hu X., Lu Y., Liu Z., Wang L. An empirical study of GPT-3 for few-shot knowledge-based VQA. 2021 URL: arXiv preprint arXiv:2109.05014.

8. Lin X., Rivenson Y., Yardimci N.T., Veli M., Luo Y., Jarrahi M., Ozcan A. All-optical machine learning using diffractive deep neural networks // Science. 2018. V. 361. № 6406. pp. 98–100.

9. Bain M., Nagrani A., Varol G., Zisserman A. Frozen in time: A joint video and image encoder for end-to-end retrieval // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. pp. 172–173.

10. Choromanski K., Likhoshesterov V., Dohan D., Song X., Gane A., Weller A. Rethinking attention with performers // 9th International Conference on Learning Representations. 2021 URL: arXiv preprint arXiv:2009.

11. Ouyang L., Wu J., Jiang X., Almeida D., Wainwright C.L., Mishkin P., Christiano P. Training language models to follow instructions with human



feedback // Advances in neural information processing systems. 2022. V. 35. pp. 27730-27744.

12. Kreiss, E., Goodman, N. D., & Potts, C. Tackling image accessibility with context: Concadia dataset. 2021. URL: arXiv preprint arXiv:2104.08376.

Дата поступления: 9.12.2024

Дата публикации: 26.01.2025