

УДК 004.048

М.Д. ДАНИЛИН, Д.А. ШЕСТОПАЛОВА
(*danilinbox2@gmail.com, shestopalova7744@gmail.com*)
Российский экономический университет им. Г.В. Плеханова

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ*

Рассматриваются потенциальные эффекты от использования больших языковых моделей в системах поддержки принятия решений и возможные варианты реализации с точки зрения архитектуры программных приложений.

Ключевые слова: *большая языковая модель, система поддержки принятия решений, LLM, СППР, архитектура СППР.*

Большая языковая модель (Large Language Model, **LLM**) – это нейронная сеть со множеством параметров (обычно миллиарды весовых коэффициентов и более), обученная на большом количестве размеченного текста с использованием обучения без учителя [3].

Большие языковые модели умеют выполнять широкий спектр операций, связанных с текстом:

1. Понимание естественного языка (Natural Language Understanding, NLU): эти модели отлично справляются с пониманием и интерпретацией человеческого языка. Они способны понимать текст на различных языках, включая разговорный, и различать контекст, тон и намерения.

2. Генерация текста: большие лингвистические модели могут генерировать человекоподобные тексты, включая статьи, рассказы, стихи и даже код. Они могут использоваться для создания контента, автоматического написания текстов и решения творческих задач.

3. Ответы на вопросы: они могут отвечать на вопросы, заданные на естественном языке, извлекая соответствующую информацию из заданного текстового источника. Это удобно для быстрого ответа на запросы пользователей.

4. Извлечение данных: большие лингвистические модели могут извлекать структурированную информацию из неструктурированного текста, например, контактные данные, даты или спецификации продуктов, что облегчает ввод и анализ данных.

5. Резюме текста: эти модели позволяют резюмировать объемные документы или статьи, преобразовывая основные моменты и ключевую информацию в более короткую и удобную для восприятия форму.

Таким образом, большие языковые модели могут быть применимы в самых разных сферах деятельности, выполняя разные задачи, связанные с текстом или данными. Доказательством эффективности работы таких сетей и применимости их на практике стал выпуск LLM GPT-3.5 компанией OpenAI в 2020 г. и дальнейший бум развития этого направления нейронных сетей. Появилось множество решений, как проприетарных, так и с открытым исходным кодом, которые можно использовать в личных и коммерческих целях.

Система поддержки принятия решений (СППР) – комплекс математических и эвристических методов и моделей, объединенных общей методикой формирования альтернатив управленческих решений в организационных системах, определения последствий реализации каждой альтернативы и обоснования выбора наиболее приемлемого управленческого решения [2].

Системы поддержки принятия решений (СППР) используются для помощи лицам, принимающим решения, в процессе анализа и обработки сложных данных. СППР помогают структурировать информацию, выявлять закономерности и предоставлять рекомендации для принятия решений в различных областях. Основные функции систем поддержки принятия решений включают:

* Работа выполнена под руководством Попова А.А., кандидата технических наук, доцента кафедры информатики ФГБОУ ВО «РЭУ им. Г.В. Плеханова».

1. Сбор и анализ данных: СППР собирают данные из различных источников, таких как базы данных, отчеты и внешние источники информации. Затем данные анализируются и представляются в виде таблиц, графиков и диаграмм, что помогает лицам, принимающим решение, получить целостное представление о ситуации.

2. Прогнозирование: СППР используют исторические данные для создания моделей, которые могут предсказывать будущие тенденции и события. Это помогает принимать решения на основе будущих сценариев и возможностей.

3. Генерация рекомендаций: СППР генерируют рекомендации и предложения на основе анализа данных и моделей. Эти рекомендации могут быть использованы для выбора оптимальных решений, управления рисками или определения стратегий для достижения определенных целей.

4. Визуализация данных: СППР предоставляют удобные и наглядные способы визуализации данных, которые облегчают понимание и интерпретацию информации.

СППР сегодня – это сложные программные продукты на основе разных технологий, работа с которыми требует от пользователя продвинутых технических навыков работы с программным обеспечением. Использование LLM в ССПР может снизить когнитивную нагрузку на лицо, принимающее решение (ЛПР), поскольку он начинает работать с естественным языком, а не с числами. Это ускоряет работу по принятию решений и может влиять на их качество. В связи с тем, что LLM являются инструментами общего назначения, которые могут быть дообучены под конкретные задачи, их применение в контексте больших систем может быть многогранным. Более того, эти технологии хорошо сочетаются друг с другом, т. к. СППР использует и генерирует большое количество данных, которые LLM может использовать как источник для последующей обработки.

Целью исследования является изучение возможности применения LLM в СППР и эффективности интеграции этой технологии. Задачами исследования являются разработка вариантов интеграции LLM и СППР, анализ сценариев использования, поиск примеров реализации такой интеграции и перечисление её основных рисков. Научная новизна нашей статьи заключается в том, что никто до этого не рассматривал применения языковых моделей как часть комплексного программного продукта СППР. В исследовании используются методы моделирования, описания, анализа и синтеза.

ССПР классифицируются по многим параметрам. По типу используемого инструментария системы СППР бывают:

- Model Driven: в основе лежат классические модели (линейные модели, модели управления запасами, транспортные, финансовые и т. п.);
- Data Driven: на основе исторических данных;
- Communication Driven: системы на основе группового принятия решений экспертами (системы фасилитации обмена мнениями и подсчета средних экспертных значений);
- Document Driven: проиндексированное (часто — многомерное) хранилище документов;
- Knowledge Driven: на основе знаний. При чем знаний как экспертных, так и выводимых машинно.

На основе базовой архитектуры СППР, описанной в статье «Интеллектуальные системы поддержки принятия решений – краткий обзор» [1], мы разработали варианты интеграции больших языковых моделей в ССПР (см. рис. на с. 68).

На этапе «Анализ предметной области» LLM может служить интерфейсом постановки пользователем задачи на естественном языке. Вместо того, чтобы самому определять путь решения проблемы и выбирать необходимые источники данных, пользователь может просто описать свою проблему, а затем модель запустит процесс исследования этой проблемы и предложит различные уточняющие вопросы, тем самым сделав запрос точнее. Более того, LLM на этом этапе может быть использована как оптимизатор запросов, который позволит эффективнее работать моделям на других этапах СППР [4].

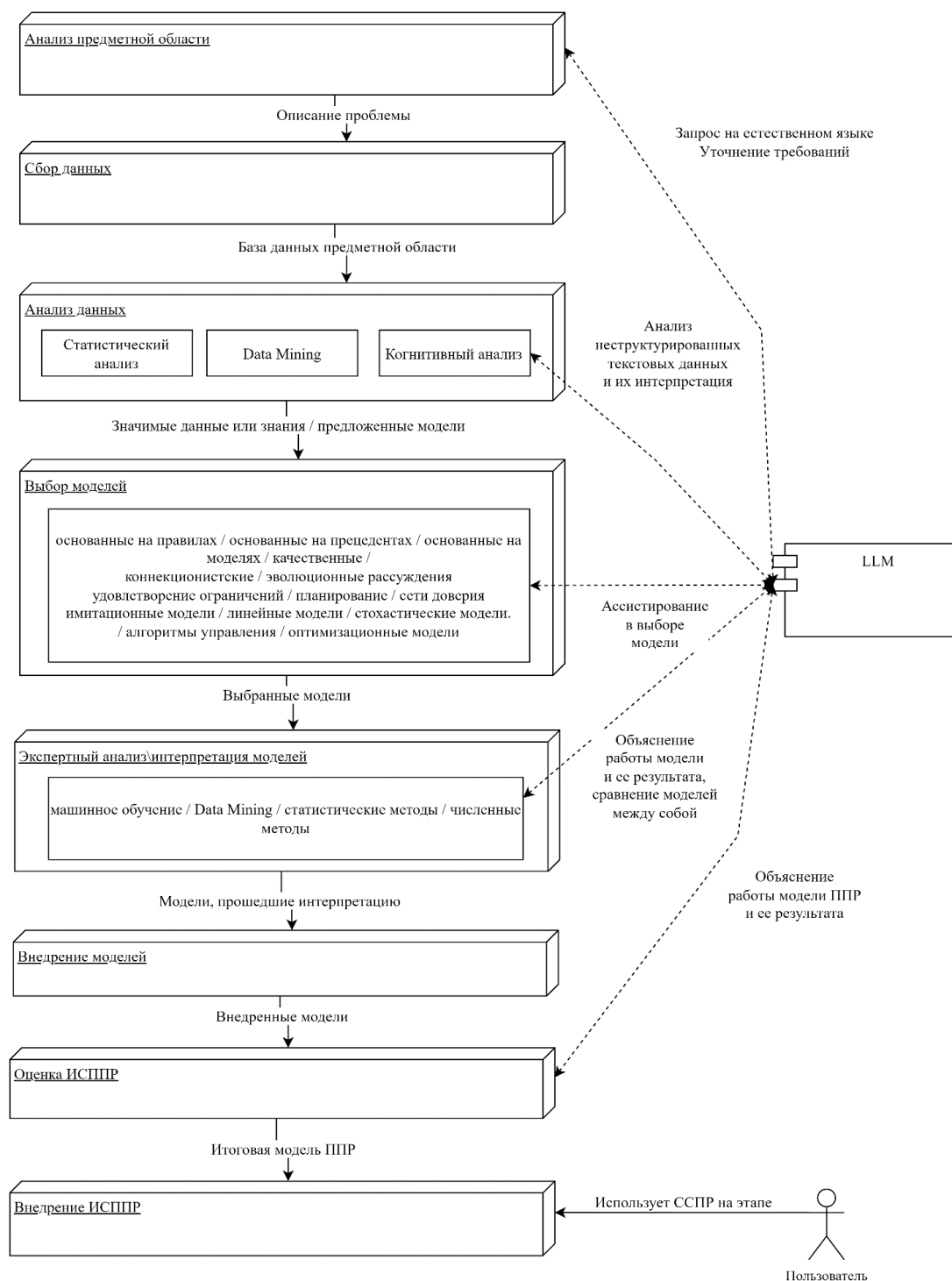


Рис. Архитектура СППР с интеграцией больших языковых моделей

На этапе «Анализ данных» большая языковая модель может помочь в обработке больших объемов текстовых данных и извлечения из них важной информации. Этот процесс является эффективнее про-

стого индексирования, т. к. значительно сокращает когнитивную нагрузку при поиске информации, за счет того, что нейросеть сама интерпретирует данные. Такую модель можно обучить на любых текстовых данных и даже медиафайлах, если использовать дополнительные модели по распознаванию речи. Например, финансовый конгломерат “Morgan Stanley” проиндексировал все документы в базе знаний компании и создал чат-бота, который на основе изученных данных помогает формировать ответы сотрудникам поддержки, предоставляет необходимую информацию для руководителей. Внедрение чат-бота многократно сократило скорость формирования ответов и принятия решений [7].

На этапе «Выбор моделей» LLM может составить краткий анализ доступных моделей и обосновать предлагаемый выбор, основываясь на контексте задачи и общих знаниях об особенностях применения тех или иных моделей. Также стоит отметить, что помощь со стороны LLM облегчит использование СППР начинающими пользователями, которые ещё не освоили модели, используемые системой в той степени, чтобы принимать обоснованные решения о применении тех или иных инструментов принятия решений.

На этапе «Интерпретация моделей» языковая модель может делать выводы по полученным итогам, облегчая пользователю процесс интерпретации результатов. Зачастую результатом работы модели является набор значений параметров (например, статистических, таких как R^2 , дисперсии и др.), которые нужно оценить в комплексе для того, чтобы сформулировать вывод. LLM (возможно специализированная именно под трактовку таких моделей) может делать это за пользователя, чем поддержит его собственный анализ или укажет на необходимость дополнительной проверки.

В результате интеграция языковых моделей происходит на всех стадиях работы ССПР, что создает новый уровень работы с системой, облегчает обработку информации, выбор подходящих средств для анализа, позволяет взаимодействовать с системой при помощи текста и мн. др.

Основная сложность при внедрении большой языковой модели в ССПР заключается в необходимости её дообучения. Если в компании используются собственные бизнес-правила или математические модели принятия решений, требуется большой объем данных, чтобы обучить модель правильно их интерпретировать. Тем не менее, разработчики СППР могут обучить LLM работе с популярными математическими и прочими моделями и использовать их в коммерческих решениях без необходимости дополнительного обучения.

Также стоит отметить, что LLM сегодня еще находятся на этапе развития и изучения и имеют характерные проблемы в корне самой технологии, которые пока не решены до конца. В их число входят:

1. Выдумывание фактов (также известное как «галлюцинации»): модели могут генерировать информацию, которая кажется правдоподобной, но на самом деле является вымышленной или неточной. Это может быть особенно проблематично при ответах на вопросы, требующие точных данных [8].

2. Деградация после обучения на искусственных данных: если модель обучается на данных, которые были искусственно сгенерированы или изменены, это может привести к ухудшению ее производительности. Модель может перенять шум или искажения из обучающих данных [5].

3. Смещение и предвзятость: языковые модели могут воспроизводить и усиливать существующие предвзятости в данных, на которых они обучались. Это может привести к нежелательным или дискриминационным выводам. В контексте ССПР эта проблема может быть актуальна при обработке данных из документов компании и использовании описания ситуации как стартовой точки [6].

4. Недостаток объяснимости: большие языковые модели часто работают как «черные ящики», что затрудняет понимание того, почему они делают определенные выводы и предсказания.

5. Проблемы с конфиденциальностью: если модель обучается на чувствительных данных, она может случайно раскрывать эту информацию в своих ответах.

6. Проблемы с надежностью: модели могут иногда давать непредсказуемые или неконсистентные ответы на одинаковые или очень похожие запросы.

По результатам исследования была разработана архитектура интеграции LLM с СППР, проанализированы возможные варианты интеграции и сделан вывод, что использование LLM в СППР возможно и является эффективным, но на данном этапе развития технологии LLM сопряжено с рисками, которые необходимо учитывать при интеграции этой технологии. Результаты исследования могут быть применены как фундамент для разработки интеграции, так и для дальнейших теоретических работ по устранению и минимизации рисков, описанных в статье.

Литература

1. Интеллектуальные системы поддержки принятия решений – краткий обзор // Хабр. [Электронный ресурс]. URL: <https://habr.com/ru/companies/ods/articles/359188/> (дата обращения: 07.10.2023).
2. Прокопенко Н.Ю. Системы поддержки принятия решений на базе Deductor Studio Academic 5.3. Н. Новгород: ННГАСУ, 2017. [Электронный ресурс]. URL: <https://bibl.nngasu.ru/electronicresources/uch-metod/programming/866785.pdf#:~:text=Система%20поддержки%20принятия%20решений%20-,выбора%20наиболее%20приемлемого%20управленческого%20решения> (дата обращения: 21.09.2023).
3. Что такое большие языковые модели? [Электронный ресурс]. URL: [https://aws.amazon.com/ru/what-is/large-language-model/#:~:text=Большие%20языковые%20модели%20\(LLM\)%20-,и%20декодера%20с%20возможностью%20самонаблюдения](https://aws.amazon.com/ru/what-is/large-language-model/#:~:text=Большие%20языковые%20модели%20(LLM)%20-,и%20декодера%20с%20возможностью%20самонаблюдения). (дата обращения: 21.09.2023).
4. Chengrun Yang, Xuezhi Wang, Yifeng Lu, Hanxiao Liu, Quoc V. Le, Denny Zhou, Xinyun Chen. Large Language Models as Optimizers // Cornell University. [Электронный ресурс]. URL: <https://arxiv.org/abs/2309.03409> (дата обращения: 21.09.2023).
5. Ilya Shumailov, Zakhar Shumaylov, Yiren Zhao, Yarin Gal, Nicolas Papernot, Ross Anderson. The Curse of Recursion: Training on Generated Data Makes Models Forget // Cornell University. [Электронный ресурс]. URL: <https://arxiv.org/abs/2305.17493> (дата обращения: 09.10.2023).
6. Large language models are biased. Can logic help save them? // Massachusetts Institute of Technology. [Электронный ресурс]. URL: <https://news.mit.edu/2023/large-language-models-are-biased-can-logic-help-save-them-0303> (дата обращения: 09.10.2023).
7. Morgan Stanley. [Электронный ресурс]. URL: <https://openai.com/customer-stories/morgan-stanley> (дата обращения: 30.09.2023).
8. Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Yejin Bang, Wenliang Dai, Andrea Madotto, Pascale Fung. Survey of Hallucination in Natural Language Generation // Cornell University. [Электронный ресурс]. URL: <https://arxiv.org/abs/2202.03629> (дата обращения: 09.10.2023).

MAKSIM DANILIN, DARYA SHESTOPALOVA

Plekhanov Russian University of Economics

THE USE OF THE LARGE LANGUAGE MODELS IN THE DECISION SUPPORT SYSTEMS

The article deals with the potential effects from the use of the large language models in the decision support systems and the potential variants of the implementation from the perspectives of the architecture of the software applications.

Key words: *large language model, decision support system, LLM, DSS, architecture of DSS.*