

УКД 33**Экономические аспекты применения технологии больших языковых моделей при анализе корпоративных данных****Меладзе Илларион Зурабович**

Магистр,
Российская академия народного хозяйства и государственной
службы при Президенте Российской Федерации,
119571, Российской Федерации, Москва, пр-кт. Вернадского, 82;
e-mail: illarionmeladze@gmail.com

Аннотация

В этой статье рассматриваются возможности и преимущества интеграции больших языковых моделей в системы корпоративной бизнес-аналитики с целью повышения доступности и удобства использования этих систем для пользователей без технической подготовки. В исследовании изучается потенциал больших языковых моделей, таких как GPT-4, для упрощения интерфейсов систем бизнес-аналитики за счет включения запросов на естественном языке, тем самым снижая потребность в специализированной технической поддержке и обучении пользователей. Благодаря анализу существующей литературы и оценке текущих ограничений систем BI, подчёркиваются финансовые выгоды, связанные с интеграцией LLM. Результаты показывают, что LLM могут значительно демократизировать доступ к данным, улучшить принятие решений на основе данных и сократить эксплуатационные расходы. Несмотря на такие проблемы, как технические сложности и сопротивление изменениям, интеграция LLM в платформы BI представляет собой многообещающий подход для организаций, стремящихся улучшить свои возможности анализа данных и общую эффективность.

Для цитирования в научных исследованиях

Меладзе И.Р. Экономические аспекты применения технологии больших языковых моделей при анализе корпоративных данных// Экономика: вчера, сегодня, завтра. 2024. Том 14. № 7А. С. 563-570.

Ключевые слова

Большие языковые модели, BI, корпоративные данные, анализ данных.

Введение

Статья посвящена использованию больших языковых моделей LLM в системах анализа данных бизнеса BI. Основное внимание уделяется повышению доступности систем анализа данных бизнеса BI для пользователей без технической подготовки, а также анализу влияния использования систем BI вместе с большими языковыми моделями LLM на затраты организации, связанные с аналитикой.

Цель статьи - исследовать, каким образом интеграция больших языковых моделей в системы анализа данных BI может улучшить доступность и удобство использования этих систем для пользователей без технической подготовки, а также оценить влияние подобных интеграций на сокращение затрат на аналитику. Результаты исследования могут привести к использованию более инклюзивных и эффективных практик анализа данных в организациях.

Анализ существующих исследований и публикаций, касающихся BI систем анализа корпоративных данных и использования больших языковых моделей LLM в различных областях. Исследование текущего состояния платформ BI и их ограничений. Оценка финансовых и производственных показателей внедрения LLM в контексте BI-платформ.

Интеграция LLM в BI-платформы значительно улучшает их доступность для пользователей без технического бэкграунда за счет упрощения интерфейсов и возможности использования естественного языка для запросов. Снижение нужды в специализированной технической поддержке и обучении ведет к уменьшению затрат, связанных с аналитикой. Ускорение процесса получения и интерпретации данных способствует повышению продуктивности сотрудников и более эффективному принятию решений.

Использование больших языковых моделей в системах BI представляет собой мощный инструмент для демократизации доступа к данным и повышения продуктивности отделов аналитики организаций. Интеграция LLM улучшает пользовательский опыт, упрощает взаимодействие с системами BI и позволяет экономить ресурсы за счет снижения нужды в специализированной поддержке. Несмотря на существующие ограничения и вызовы, такие как технические трудности и организационное сопротивление, преимущества использования LLM в BI значительно перевешивают возможные риски.

Инструменты и платформы бизнес-аналитики (BI) стали неотъемлемыми элементами стратегического арсенала современных предприятий, предоставляя возможность использовать данные для принятия обоснованных решений. Шестьдесят семь процентов мировой рабочей силы имеют доступ к инструментам бизнес-аналитики, а 75 процентов имеют доступ к программному обеспечению для анализа данных [Аксенчер. Влияние грамотности в области данных на человека 2020 г., www...].

Однако традиционно использование инструментов BI было ограничено сферой технически квалифицированных аналитиков и специалистов по обработке данных, что усложняло задачу полного использования потенциала аналитики, основанной на данных [Бареддо, Диас, Дель, 2020].

Платформы Self-Service BI были разработаны как решение для демократизации доступа к данным и расширения возможностей пользователей без технической подготовки. Эти платформы направлены на снижение зависимости от специализированной ИТ-поддержки, предоставляя интуитивно понятные интерфейсы, которые позволяют пользователям самостоятельно создавать отчеты, анализировать данные и делать выводы [Божич, 2024]. Однако, несмотря на достигнутый прогресс, остаются значительные препятствия. Пользователи без технической подготовки часто сталкиваются с препятствиями при работе со сложными

системами бизнес-аналитики BI. Затруднения вызывает интерпретация визуализаций данных и составление корректных запросов к системе, что ограничивает способность извлекать из данных значимую для анализа информацию [Джьюти, Шумбел, 2023].

Большие языковые модели (LLM) являются существенным прорывом в сфере интерпретации естественного языка. Они показывают высокую точность в анализе и интерпретации текста, что способствует более интуитивному взаимодействию с технологиями. Считается, что интеграция больших языковых моделей в платформы Self-Service BI может позволить организациям устранить оставшиеся барьеры для пользователей без технической подготовки. Многие считают, что эти модели могут упростить пользовательские интерфейсы, обеспечить выполнение запросов на естественном языке и предоставить контекстные объяснения, делая инструменты бизнес-аналитики более доступными и удобными для пользователей [Джьюти, Шумбел, 2023].

Основные трудности пользователей без технической подготовки в системах BI

Большинство актуальных систем BI требует наличие технических знаний и специализированных навыков, что является препятствием для многих пользователей. Перед пользователями встаёт потребность в изучении специальных языков запросов, таких как SQL, для выполнения аналитических задач. В результате чего пользователи имеют ограниченный доступ к различным источникам данных без помощи IT-отдела.

Время, затрачиваемое на обучение работе с системами и приобретение необходимых навыков, является существенным фактором при внедрении и использовании BI-решений. Традиционные BI-системы требуют значительных инвестиций в обучение персонала. Это приводит к временным и финансовым затратам, которые необходимо учитывать при оценке эффективности внедрения BI-решений.

Также формирование нестандартных отчётов зачастую требует значительных навыков и времени. Если пользователь захочет посмотреть на данные в каком-то новом разрезе, который до этого не использовался в компании, зачастую ему нужно привлечение аналитиков и сотрудников IT-отдела.

Помимо этого, стоит проблема с интеграцией разных источников данных в одну систему. Сотрудник, работающий с данными из различных систем (CRM, ERP, другие системы учета), может тратить значительное время на ручную интеграцию и очистку данных, что снижает эффективность и скорость анализа.

Большие языковые модели (LLM) и их понимание

Большие языковые модели (LLM) являются разновидностью искусственного интеллекта, обучаемого на больших объемах текстовых данных. В качестве примера можно привести самую известную большую языковую модель: GPT-4 (Generative Pre-trained Transformer 4), созданную компанией OpenAI. Модели такого типа могут выполнять целый спектр разных задач. Например, обрабатывать и генерировать тексты, написанные на естественном языке.

Большие языковые модели способны понимать и интерпретировать лингвистические структуры и контексты [Кьярелло, Джордано, Спада, Барандони, Фантони, 2024], что делает их полезными для различных корпоративных приложений, включая BI. Вместо использования сложных SQL-запросов, пользователи могут задавать вопросы на естественном языке и

получать точные ответы. Например, можно спросить: "Какие были наши продажи в прошлом квартале по регионам?" или "Какова была динамика продаж в прошлом месяце?" вместо составления сложного запроса.

Эти модели могут автоматизировать рутинные аналитические задачи, такие как построение отчетов, подготовка данных и даже интерпретация результатов, что освобождает время пользователей для более креативных и стратегических задач. Также эти модели могут генерировать контент, который неотличим от написанного человеком. LLM могут автоматически создавать подробные аналитические отчеты, базируясь на входных данных. Это позволяет сократить время, затрачиваемое на подготовку отчетности, и минимизировать ошибки [Мэннинг, Кларк, Хьюитт, Ханделвал, Леви, 2020].

В контексте бизнес-аналитики, LLM предлагают множество преимуществ, которые могут способствовать увеличению доступности и продуктивности BI-систем:

- Упрощение пользовательских интерфейсов. Интеграция LLM позволяет создавать более интуитивные и удобные для пользователя интерфейсы. Пользователи могут использовать естественный язык для взаимодействия с системой, избегая необходимости изучать сложные технические аспекты [Мэннинг, Кларк, Хьюитт, Ханделвал, Леви, 2020].
- Контекстуальная Помощь. Предоставление контекстуальных подсказок и объяснений помогает пользователям лучше понимать данные и делать более обоснованные выводы [Палыс, 2020].
- Автоматическая Подготовка и Очистка Данных. LLM могут автоматически идентифицировать и исправлять ошибки в данных, а также подготавливать данные для анализа без вмешательства пользователя [Партнеры NewVantage, компания Wavestone, опрос руководителей в области данных и искусственного интеллекта 2022 г., www...].
- Генерация Отчетов и Визуализаций. На основе заданных вопросов и данных модели могут создавать отчеты и визуализации, которые наилучшим образом представляют нужную информацию. Пользователь может легко корректировать и уточнять данные отчеты, уточняя свои запросы на естественном языке [Хоссейнзаде, Ажир, Ахмед, Гафур, Ахмед, Рахмани, 2021].
- Аналитические Подсказки и Объяснения. Система может давать рекомендации о том, какие метрики использовать, объяснять полученные результаты и предлагать действия на основе анализируемых данных [Хоссейнзаде, Ажир, Ахмед, Гафур, Ахмед, Рахмани, 2021].

Текущее состояние систем аналитики данных бизнеса (BI)

BI системы – это системы аналитики данных бизнеса. Они преобразуют большие объемы данных в значимую и полезную информацию, позволяя организациям принимать информированные решения и формировать стратегические планы.

Self-Service подход в системах BI — это подход, позволяющий пользователям без глубоких технических знаний самостоятельно разрабатывать и анализировать отчеты, дашборды и другие аналитические данные [Божич, 2024].

Этот подход направлен на предоставление доступа к данным и аналитическим инструментам всем уровням организации, уменьшая зависимость от ИТ-отделов и специализированных аналитиков. Упрощение доступа к данным для широкой аудитории помогает организации стать более информированной и увеличить количество решений,

принятых на основе данных [Барредо, 2020].

Таким образом упрощается и ускоряется процесс анализа данных, что позволяет быстрее принимать стратегически важные решения. Разгрузка ИТ-ресурсов от рутинных задач создания отчетов и анализа данных позволяет им концентрироваться на более сложных задачах. Современные платформы BI предлагают ряд ключевых функций, направленных на то, чтобы сделать процесс анализа данных более интуитивным и доступным.

Внедрение больших языковых моделей (LLM) сейчас можно назвать передовым направлением в развитии Self-Service подхода в системах BI.

Анализ финансовых затрат

Потенциально внедрение больших языковых моделей в системы бизнес-аналитики может предоставлять существенную экономию средств. Но также надо учитывать, что подобное внедрение требует и дополнительных расходов со стороны компании.

К расходам, которые, в основном, являются связаны с интеграцией LLM, можно отнести непосредственно саму разработку необходимого ПО, интеграцию по API и обучение персонала.

Рекуррентные расходы включают в себя плату за использование API, расходы, связанные с обслуживанием системы, плату за использование облачных сервисов, например, AWS или Google Cloud.

К рекуррентным расходам также можно отнести плату за техническую поддержку, которую взимают компании-поставщики ПО за обеспечение оперативной поддержки в случае непредвиденных ситуаций и для поддержания беспереывной работы системы.

Помимо расходов выше, при внедрении системы требуются разовые траты, связанные непосредственно с самой интеграцией сервисов LLM с текущими BI-системами и другими корпоративными программами, и приложениями.

Потенциально, экономия средств исходит из возможности снижения затрат на штатных аналитиков данных, сотрудников ИТ-поддержки. Системы BI с интегрированным ИИ способны заметно снизить нагрузку на специалистов по работе с данными, что потенциально приводит к экономии затрат на персонал [Барредо, 2020].

В зависимости от типа и размера компании, основные показатели рентабельности инвестиций могут отличаться, но в среднем, компании, которые внедрили систему BI совместно с интегрированной языковой моделью, могут ожидать рентабельность инвестиций в течение 14 месяцев с момента начала использования системы [Шакил, Аль-Акраби, Хилл, 2020].

Существенные затраты на первичную интеграцию митигируются высокой рентабельностью. Долгосрочный потенциал экономии средств показывает, что подобное решение является эффективным для организаций, стратегические решения которых во многом зависят от аналитики данных.

Выше мы рассмотрели явные финансовые преимущества использования систем BI в связке с большими языковыми моделями. Помимо них, также, есть множество других плюсов подобного подхода, как явно измеряемых, так и на первый взгляд незаметных. Автоматизация создания отчетов и аналитики компании может помочь сотрудникам меньше времени тратить на рекуррентные задачи, таким образом освобождая время и ресурсы на более стратегически важные для компании функции. Также можно заметить, что подобный подход приводит к снижению количества ошибок, связанных с человеческим фактором. Данные аналитики становятся более надёжными.

Интеграция LLM в системы BI: организация и рекомендации

Процесс интеграции больших языковых моделей в системы BI требует тщательного планирования. Подобные системы рассчитываются на многолетнее использование, поэтому для качественной реализации необходим всесторонний анализ требующихся функций системы и требований пользователей к аналитике и отчётности. Помимо этого, на подобных проектах внедрения всегда требуется создание обучающих материалов для пользователей и проведение тренингов для того, чтобы пользователи максимально быстро адаптировались к новым системам и максимально эффективно начали использовать их функции.

Когда система BI с интегрированной большой языковой моделью будет успешно внедрена, компании могут понадобится метрики и процессы для оценки работы системы и для мониторинга её производительности. В первую очередь с этим может помочь сбор обратной связи пользователей для поиска узких мест в системе и внесения возможных улучшений.

Для того, чтобы оценить эффективность системы BI, можно использовать следующие метрики:

- Скорость получения аналитики. К этой метрике можно отнести временные затраты в часах на создание пользователем отчетов и аналитики.
- Сбор обратной связи пользователей. Можно использовать анкеты и опросы для измерения удовлетворённости пользователей функциями системы.
- Аналитика затрат. Подсчёт коэффициента рентабельности поможет понять, за какое время вернулись инвестиции на внедрение.
- Подсчёт количества ошибок в аналитике. Измерение количества ошибок в аналитике поможет убедиться в правильности работы системы и её функций.

Заключение

В этой статье мы исследовали, как интеграция в системы корпоративной аналитики BI больших языковых моделей способна повлиять на удобство использования аналитических систем пользователями без технической подготовки и без обширного обучения пользователей анализу данных. Также мы оценили влияние подобных интеграций на сокращение затрат на аналитику.

Большие языковые модели являются мощным инструментом, позволяющим пользователям улучшить эффективность работы в имеющихся BI системах. Системы становятся более удобны и доступны для пользователя без технической подготовки.

Интеграция больших языковых моделей требует инвестиций ресурсов на внедрение и обучение пользователей, но мы выяснили, что при правильном подходе к планированию, подобные инвестиции способны довольно быстро окупиться, начать помогать компании адаптироваться к изменчивым рыночным условиям и повысить конкурентоспособность.

Библиография

1. Аксенчер. Влияние грамотности в области данных на человека. Руководство для лидеров по демократизации данных, повышению производительности и расширению прав и возможностей сотрудников. 2020. С. 5. URL: <https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-118/accenture-the-human-impact-data-literacy.pdf> (дата обращения: 13.10.2020).
2. Барредо А.А., Диас – Родригес Н., Дель Сер Х. и др. Объяснимый искусственный интеллект (XAI): концепции, таксономии, возможности и проблемы к ответственному ИИ. Information Fusion. 2020. Том 58. С. 82-115.

3. Божич В. Влияние искусственного интеллекта на бизнес-аналитику. 2024. URL: https://www.researchgate.net/publication/377118301_The_Impact_of_Artificial_Intelligence_on_Business_Intelligence (дата обращения: 11.01.2024).
4. Джьоти Р., Шубмел Д. Бизнес-возможности искусственного интеллекта. 2023. The Business Opportunity of AI. URL: https://idcdocserv.com/US51315823-IG-ADA?trk=article-ssr-frontend-pulse_little-text-block (дата обращения: 08.11.2024).
5. Кьярелло Ф., Джордано В., Спада И., Барандони С., Фантони Г. Будущие приложения генеративных больших языковых моделей: исследование на основе данных в ChatGPT. Technovation. 2024. Т. 133. 103002. С. 0166-4972. URL: <https://doi.org/10.1016/j.technovation.2024.103002>. (дата обращения: 29.03.2024).
6. Мэннинг К.Д., Кларк К., Хьюитт Дж., Ханделвал У., Леви О. Возникающая лингвистическая структура в искусственных нейронных сетях, обученных самоконтролем. Труды Национальной академии наук. 2020. 117(48), С. 30046-30054.
7. Палыс М., Палыс А. Преимущества и проблемы внедрения самостоятельной бизнес-аналитики. Procedia Computer Science. Т. 225. 2023. С. 795-803.
8. Партнеры New Vantage, компания Wavestone, опрос руководителей в области данных и искусственного интеллекта. Data and AI Leadership Executive Survey. 2022. С. 7. URL: https://c6abb8db-514c4f5bb5a1fc710f1e464e.filesusr.com/ugd/e5361a_2f859f3457f2_4cff9b2f8a2bf54f82b7.pdf (дата обращения: 22.01.2022).
9. Хоссейнзаде М., Ажир Е., Ахмед О., Гафур М., Ахмед С., Рахмани А., Во Б. Механизмы очистки данных и подходы для аналитики больших данных: систематическое исследование. Журнал Ambient Intelligence и гуманизированные вычисления. 2021. № 14. С. 1-13.
10. Шакил Х.М., и Ирам Ш., Аль-Акраби Х., Аль-Акраби А., Хилл Р. Комплексный обзор современного состояния инструментов визуализации данных: исследовательские разработки, проблемы и будущая структура визуализации для конкретных областей. 2020. С. 96581-96601.
11. Шен Л. и др. На пути к интерфейсам на естественном языке для визуализации данных: обзор. 2022. Т. 20. №. 10 (20). URL: https://www.researchgate.net/publication/354460810_Towards_Natural_Language_Interfaces_for_Data_Visualization_A_Survey обзор (дата обращения: 08.09.2021).

Economic aspects of using large language models technology in corporate data analysis

Illarion Z. Meladze

Master of Economics,
Russian Presidential Academy of National Economy and Public Administration,
119571, 82, Vernadsky Ave., Moscow, Russian Federation;
e-mail: illarionmeladze@gmail.com

Abstract

In this article we will explore the opportunities and benefits of integrating large language models into enterprise business intelligence systems to improve the accessibility and usability of these systems for non-technical users. The study examines the potential of LLMs, such as GPT-4, to simplify BI interfaces by enabling natural language queries, thereby reducing the need for specialized technical support and training. Through an analysis of existing literature and evaluation of current BI system limitations, the research highlights the financial benefits associated with LLM integration. The findings suggest that LLMs can significantly democratize data access, improve data-driven decision-making, and reduce operational costs. Despite challenges such as technical complexities and resistance to change, the integration of LLMs into BI platforms presents a promising approach for organizations aiming to enhance their data analysis capabilities and overall productivity.

For citation

Meladze I.Z. (2024) Ekonomicheskiye aspekty primeneniya tekhnologii bol'shikh yazykovykh modeley pri analize korporativnykh dannykh [Economic aspects of using large language models technology in corporate data analysis]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 14 (7A), pp. 563-570.

Keywords

Large Language Models, BI, enterprise data, data analysis.

References

1. Accenture. (2020) Vliyaniye gramotnosti v oblasti dannykh na cheloveka. Rukovodstvo dlya liderov po demokratizatsii dannykh, povysheniyu proizvoditel'nosti i rasshireniyu prav i vozmozhnostey sotrudnikov [The Human Impact of Data Literacy: A Leader's Guide to Democratizing Data, Improving Productivity, and Empowering Employees] p. 5. URL: <https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-118/accenture-the-human-impact-data-literacy.pdf> (accessed: 13.10.2020).
2. Bozic V. (2024) Vliyaniye iskusstvennogo intellekta na biznes-analitiku [The Impact of Artificial Intelligence on Business Intelligence] URL: https://www.researchgate.net/publication/377118301_The_Impact_of_Artificial_Intelligence_on_Business_Intelligence (accessed: 11.01.2024).
3. Barredo AA, Diaz-Rodriguez N., Del Ser J. et al. (2020) Ob"yasnimyy iskusstvennyy intellekt (XAI): kontseptsii, taksonomii, vozmozhnosti i problemy k otvetstvennomu II. [Information Fusion] Vol. 58. pp. 82-115.
4. Chiarello F., Giordano V., Spada I., Barandoni S., Fantoni G. (2024) Budushchiye prilozheniya generativnykh bol'shikh yazykovykh modeley: issledovaniye na osnove dannykh v ChatGPT [Technovation] Vol. 133. 103002, pp. 0166-4972. URL: <https://doi.org/10.1016/j.technovation.2024.103002>. (Accessed: 29.03.2024).
5. Hosseinzadeh M, Azhir E, Ahmed O, Ghafour M, Ahmed S, Rahmani A., Vo B. (2021) Mekhanizmy ochistki dannykh i podkhody dlya analitiki bol'shikh dannykh: sistematicheskoye issledovaniye [Journal of Ambient Intelligence and Humanized Computing] №. 14. pp. 1-13.
6. Jyoti R., Shubmel D. (2023) Biznes-vozmozhnosti iskusstvennogo intellekta [The Business Opportunity of AI] URL: https://idcdocserv.com/US51315823-IG-ADA?trk=article-ssr-frontend-pulse_little-text-block (accessed: 08.11.2024).
7. Manning K.D., Clark K., Hewitt J., Khandelwal U., Levy O. (2020) Voznikayushchaya lingvisticheskaya struktura v iskusstvennykh neyronnykh setyakh, obuchennykh samokontrolem [Proceedings of the National Academy of Sciences] 117(48). pp. 30046-30054.
8. Partnery NewVantage, kompaniya Wavestone, opros rukovoditeley v oblasti dannykh i iskusstvennogo intellekta. (2022) [Data and AI Leadership Executive Survey] p. URL: https://c6abb8db-514c4f5bb5a1fc710f1e464e.filesusr.com/ugd/e5361a_2_f859f3457f2_4cff9b2f8a2bf54f82b7.pdf (accessed: 22.01.2022).
9. Palys M. (2023) Preimushchestva i problemy vnedreniya samostoyatel'noy biznes-analitiki Procedia [Computer Science] Vol. 225. p.795-803.
10. Shakeel H.M., Iram, S., Al-Akrabi, H., Al-Akrabi, A., and Hill, R. (2020) Kompleksnyy obzor sovremennogo sostoyaniya instrumentov vizualizatsii dannykh: issledovatel'skiye razrabotki, problemy i budushchaya struktura vizualizatsii dlya konkretnykh oblastey [Comprehensive Review of the Current State of Data Visualization Tools: Research Developments, Challenges, and Future Domain-Specific Visualization Frameworks] pp. 96581-96601.
11. Shen L. et al. (2022) Na puti k interfeysam na yestestvennom yazyke dlya vizualizatsii dannykh: obzor [Towards natural language interfaces for data visualization: a review] Vol. 20. № 10 (20). URL: https://www.researchgate.net/publication/354460810_Towards_Natural_Language_Interfaces_for_Data_Visualization_A_Surveyreview (accessed: 09.08.2021).