

ПРОЕКТИРОВАНИЕ АРХИТЕКТУРЫ ML-СИСТЕМЫ ФОРМИРОВАНИЯ НОВОСТНЫХ ДАЙДЖЕСТОВ

Солодкая М.А. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Кугаевских А.В.
(ИТМО)

Введение. В условиях стремительного роста объемов информации в интернете пользователям становится все сложнее анализировать актуальные события и выявлять ключевые взаимосвязи между ними. Традиционные методы агрегирования новостей зачастую ограничиваются простым сбором заголовков и кратких аннотаций, не учитывая скрытые взаимосвязи между событиями, компаниями и личностями. Однако, понимание таких связей критически важно для аналитиков, журналистов и исследователей, позволяя выявлять тренды, предсказывать развитие событий и получать более полное представление о происходящем.

Решением этой проблемы может стать разработка ML-системы, способной анализировать различные источники информации, выявлять связи между сущностями, строить графы взаимосвязей и формировать детализированные дайджесты с объяснениями скрытых и явных взаимосвязей. Такая система позволит автоматизировать процесс обработки новостей, обеспечивая пользователей глубокой аналитикой на основе машинного обучения и обработки естественного языка (NLP).

Основная часть. Для разработки ML-системы формирования новостных дайджестов необходимо спроектировать её архитектуру. Разработка архитектуры ML-системы формирования новостных дайджестов предполагает использование больших языковых моделей (LLM) для обработки пользовательского запроса и автоматического сбора информации из открытых источников. Такими моделями могут быть DeepSeek или Qwen. Оба семейства моделей можно использовать локально, обучить на своём наборе данных и исследовать их применимость в поставленной задаче [1,2]. Процесс работы системы предполагает несколько ключевых этапов. После получения пользовательского запроса LLM инициирует сбор релевантных источников из интернета, используя API новостных агрегаторов, поисковые системы и базы данных. Тексты статей проходят предобработку: удаление дубликатов, нормализация, фильтрация нерелевантного контента. Затем с помощью LLM происходит генерация summary (краткое содержание) и извлечение метаданных из новостных материалов, выделяя ключевые события, тенденции и взаимосвязи. Из текста извлекаются различные метаданные. Собранные источники и метаданные поступают в графовую нейронную сеть, которая выполняет определение существующих связей между объектами (например, кто с кем взаимодействовал, какие компании упоминаются в одном контексте и т. д.), а также предсказание возможных новых взаимосвязей на основе анализа графа, учитывая закономерности в данных. На основе анализа строится граф взаимосвязей, позволяющий пользователю наглядно видеть структуру новостного контекста [2]. Визуализация включает фильтрацию и настройку отображаемых взаимосвязей, позволяя детально исследовать ключевые узлы.

Выводы. Результаты этого исследования могут быть использованы для разработки ML-системы, которая позволит автоматизировать процесс анализа новостей, обеспечивая пользователей детализированными дайджестами с выявленными взаимосвязями. Это откроет новые возможности в аналитике новостей, обеспечивая глубокий контекстный анализ событий и персон.

Список использованных источников:

1. Jinze Bai, Shuai Bai & oth. (2023). Qwen Technical Report [Электронный ресурс]. URL: <https://arxiv.org/abs/2309.16609> (дата обращения: 10.01.2025).
2. Быстрое локальное развертывание DeepSeek, Рыбаков Алексей. 2025. [Электронный

ресурс]. URL: <https://habr.com/ru/articles/879076/> (дата обращения: 05.02.2024).

3. Hamilton, W. L., Ying, R., Leskovec, J. (2017). Representation Learning on Graphs: Methods and Applications. [Электронный ресурс]. URL: <https://arxiv.org/abs/1709.05584> (дата обращения: 26.01.2025).