

УДК 004.891.3

ПРИМЕНЕНИЕ АРХИТЕКТУРЫ RETRIEVAL AUGMENTED GENERATION ДЛЯ ФОРМИРОВАНИЯ ИНТЕРПРЕТИРУЕМЫХ ДИАГНОСТИЧЕСКИХ ЗАКЛЮЧЕНИЙ В ЦИТОЛОГИИ ЩИТОВИДНОЙ ЖЕЛЕЗЫ ПО СИСТЕМЕ BETHESDA

Основин С.С. (НИЯУ МИФИ), Садохин А.А. (НИЯУ МИФИ)

Научный руководитель – доктор технических наук, профессор Зайцев К.С. (НИЯУ МИФИ)

Введение. В настоящий момент системы искусственного интеллекта, основанные на больших языковых моделях, демонстрируют крайне высокую эффективность в задачах анализа и интерпретации различных биомедицинских данных, включая визуальную оценку цитологических изображений. Однако, признавая риски и различные ограничения из-за сложной и непрозрачной системы диагностических решений их внедрение в клиническую практику может быть ограничено поскольку медицинские работники не могут полноценно доверять подобным системам. В условиях задачи медицинской диагностики, когда ошибки могут напрямую влиять на то, какой способ лечения будет выбран, например, оперативное вмешательство или периодическое наблюдение. Критически важна возможность объяснения причинно-следственных связей, лежащих в основе поставленного диагноза. Данная проблема особенно актуальна для цитологической диагностики щитовидной железы, где интерпретация морфологических признаков требует строгого соответствия принципам и правилам. Этого позволяет достичь международная система классификации Bethesda [1]. Задачей настоящего исследования является разработка специализированной системы помощника для медицинского работника, способной генерировать не только диагностическую категорию, но и клиническую интерпретацию или обоснование решения на основе интеграции количественных признаков изображений и структурированного экспертного знания.

Основная часть. Исследование реализовано в виде модульной системы, использующего в своей работе два независимых источника данных:

- 1) Сформированный и предобработанный корпус знаний на основе официального руководства TBSRTC (Системы Bethesda с описаниями категорий, параметрическими критериями и значения риска злокачественности с рекомендациями для врача) [2];
- 2) Количественные признаки, извлечённые реализованной смежной командой внешней системы визуальной сегментации из цитологических изображений (количество и типы клеточных структур, площадь ядра, соотношение ядро/цитоплазма, текстурные характеристики).

Архитектура системы включает следующие компоненты:

- 1) Модуль семантического поиска (эмбединговая модель thenlper/gte-small, индексация в FAISS) для извлечения релевантных фрагментов из корпуса TBSRTC;
- 2) Модуль нормализации числовых признаков, полученных от сегментационного инструмента;
- 3) Модуль формирования контекстного промпта, строго привязывающего генерацию к извлечённым данным [3];
- 4) Генеративное ядро на базе открытых LLM (тестирование проведено на Qwen2.5-7B-Instruct, Llama-3.1-8B-Instruct, Gemini 1.5 Pro, GigaChat-20B-A3B-Instruct, DeepSeek-LLM-7B-Chat.)

Проверка и валидация системы выполнена на выборке обезличенных клинических случаев, предоставленных НМИЦ эндокринологии им. академика И.И. Дедова Минздрава России. Помимо этого, тестирование проводилось с использованием стандартных метрик для определения количество правильных постановок меток Bethesda использовались F1-Score, Recall и Accuracy. Для корректности работы системы RAG и генеративных моделей использовались так же стандартные метрики ROUGE, Perplexity, Context Relevance, Faithfulness [4]. Результаты демонстрируют способность системы генерировать синтаксически корректные и семантически согласованные объяснения, содержащие: явное указание

диагностической категории Bethesda, перечисление доминирующих количественных признаков в цифровом выражении, их интерпретацию через призму критериев TBSRTC и логический вывод, соответствующий клинической терминологии. Ключевым достижением является снижение галлюцинаций моделей: если не использовать модуль для извлечения контекста доля некорректных ответов моделей с количеством параметров менее 7B резко возрастает. Подтверждена гипотеза о том, что механизм динамического контекстного дополнения существенно повышает фактологическую точность и доменную согласованность генерируемых объяснений даже при использовании базовых конфигураций LLM.

Выводы. В ходе проведенного исследования и разработки была продемонстрирована принципиальная возможность применения архитектуры Retrieval Augmented Generation для моделирования диагностического рассуждения в узкоспециализированной медицинской области, такой как цитология щитовидной железы. Использование количественных признаков изображений с формализованным экспертным знанием, полученным через механизм семантического поиска в векторной базе знаний, позволяет генерировать гораздо более клинически точные и понятные объяснения, соответствующие ожидаемым рассуждениям. Наше исследование показывает, насколько важно использование контекста при работе с большими языковыми моделями его использование существенно снижает частоту галлюцинаций в моделях с ограниченным количеством параметров, что делает предложенный подход применимым без ресурсоёмкого дообучения на узкоспециализированных данных. Данная система призвана сделать работу уже существующих моделей, заточенных под определенные задачи, более прозрачными и понятными для медицинского специалиста, что позволяет ускорить интерпретацию результатов, а также в следствии принять правильное решение. В дальнейших исследованиях планируется расширить подход подбора текстовых фрагментов в использовании RAG, использовать различные методы дополнения контекста и сравнить их потенциальную эффективность. Помимо этого, предложенный подход обладает потенциалом масштабирования на другие формализованные медицинские домены, где так же существуют различные диагностическими критерии и количественные оценки признаков.

Список использованных источников:

1. Ali S., Cibas E. The Bethesda system for reporting thyroid cytopathology / ed.: Ali S.Z., Cibas E.S. — Cham: Springer International Publishing, 2018. — DOI: 10.1007/978-3-319-60570-8.
2. Ali S.Z., Baloch Z.W., Cochand-Priollet B., Schmitt F.C., Vielh P., VanderLaan P.A. The 2023 Bethesda system for reporting thyroid cytopathology // *Thyroid®*. — 2023. — № 7. — DOI: 10.1089/thy.2023.0141.
3. Gupta S., Ranjan R., Singh S.N. A comprehensive survey of retrieval-augmented generation (RAG): evolution, current landscape and future directions // *arXiv preprint arXiv:2410.12837*. — 2024.
4. Gu B., Desai R.J., Lin K.J. et al. Probabilistic medical predictions of large language models // *npj Digital Medicine*. — 2024. — Vol. 7, No. 367. — DOI: 10.1038/s41746-024-01366-4.