

**Мультиагентные биомедицинские трансформенные решения проблем
щитовидной железы**

Научный руководитель – Зайцев Константин Сергеевич

Дюльдин Евгений Владимирович

Аспирант

Национальный исследовательский ядерный университет «МИФИ», Факультет
кибернетики и информационной безопасности, Москва, Россия

E-mail: Zhecos1@yandex.ru

В настоящее время для решения проблем в биомедицинском домене используются сложные системы на основе множества решателей и экспертного врачебного опыта для получения достоверных результатов. Современные модели на основе трансформенных подходов [1, 2, 3] выдают правдоподобные ответы, но финальная валидация остаётся дополнительной нагрузкой врача, для уменьшения временных затрат предлагается создание группы AI-агентов, которые не будут пытаться полностью решить предложенную биомедицинскую задачу, но автоматизируют отдельные блоки финального конвейерного решения.

Предложенное решение состоит из нескольких последовательных этапов, каждый из которых включает логические микросервисные блоки для достижения не только наиболее точных прогнозных значений, но и получения достаточного быстрого действия системы в асинхронном режиме работе с пользователями.

Модульная структура состоит из трёх этапов, где основой является мультиагентный модуль на стыке технологий NLP и ASR, что позволяет распознавать и возвращать в потоковом режиме полученные ответы, для решения задач Text to Text используются предобученные на задачу Bethesda [4] модели LLAMA [2], в комплексе с BART подобными выпрямителями выходных последовательностей. Финальный выход обрабатывается сервисом Gradio [5] и может быть развёрнут, как полноценное решение на промышленном сервере на основе FastApi[6].

Полученное решение ускоряет наиболее сложные рутинные этапы по формированию врачебного описания и создания предсказательного шаблона врачебного заключения. В результате полученный конвейер работает, как ассистент для опытного врача и помощник для начинающего специалиста.

Источники и литература

- 1) Guo D. et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning //arXiv preprint arXiv:2501.12948. – 2025.
- 2) Touvron H. et al. Llama: Open and efficient foundation language models //arXiv preprint arXiv:2302.13971. – 2023.
- 3) Jiang F. Identifying and mitigating vulnerabilities in llm-integrated applications : дис. – University of Washington, 2024.
- 4) Ali S, Cibas E. The Bethesda System for Reporting Thyroid Cytopathology. (Ali SZ, Cibas ES, eds.). Cham: Springer International Publishing; 2018. doi: <https://doi.org/10.1007/978-3-319-60570-8>
- 5) Abid A. et al. Gradio: Hassle-free sharing and testing of ml models in the wild //arXiv preprint arXiv:1906.02569. – 2019.

- 6) Howard J., Gugger S. Fastai: a layered API for deep learning // Information. – 2020. – Т. 11. – №. 2. – С. 108.