

ЭМБЕДДИНГОВЫЕ МОДЕЛИ ДЛЯ ДИАГНОСТИКИ ПАЦИЕНТОВ В УЧРЕЖДЕНИЯХ ЗДРАВООХРАНЕНИЯ

Рассматриваются перспективы применения мультимодальной эмбединговой модели Gemini Embedding 2.0 (Google) в учреждениях здравоохранения для повышения качества диагностики. Проанализированы технические характеристики модели, сценарии клинического применения и барьеры внедрения.

Введение

Нарастающий объём клинических данных – медицинских карт, результатов визуализации, аудиозаписей приёмов, видеоматериалов процедур – создаёт потребность в инструментах семантического поиска и сопоставления разнородной информации. Традиционный поиск в медицинских системах основан на ключевых словах, что ограничивает полноту и релевантность результатов. Эмбединговые модели преобразуют данные произвольной природы в числовые векторы в едином семантическом пространстве, где близость векторов отражает смысловую близость объектов. В марте 2026 года Google представила Gemini Embedding 2.0 – мультимодальную модель, принимающую текст, изображения, PDF, аудио и видео и отображающую их в 3072-мерное векторное пространство. Мультимодальность клинической информации делает данную модель перспективной для медицинской диагностики.

I. Технические характеристики и сценарии клинического применения

В контексте мультимодальности модель принимает изображения (PNG, JPEG – до 6 на запрос), PDF (до 6 страниц), аудио (MP3, WAV – до 80 с), видео (MP4, MP4 – до 120 с) и текст (до 8192 токенов). Рентгенограмма, заключение врача и аудиозапись консультации преобразуются в сопоставимые векторы в едином пространстве. 3072-мерный вектор по умолчанию обеспечивает высокую гранулярность представления. Поддержка MRL (Matryoshka Representation Learning) через параметр `output_dimensionality` позволяет уменьшить размерность, оптимизируя баланс точности и стоимости хранения. Указание контекста задачи оптимизирует эмбединги под конкретный сценарий. Встроенный OCR (Optical Character Recognition) обрабатывает отсканированные документы без предварительной оцифровки. Извлечение аудиодорожек из видео перспективно для анализа эндоскопических записей с комментариями врача. Сценарии применения: семантический кросс-модальный поиск по архиву пациента; поиск аналогичных случаев для дифференциальной диагностики; кластеризация ко-

дорт для стратификации рисков; контроль качества заключений через выявление расхождений между изображением и его описанием.

II. Барьеры внедрения

Нормативные ограничения. Передача медицинских данных в облачное хранилище Google ставит вопрос соответствия закона РБ 99-3, HIPAA, GDPR. Решение – анонимизация (обезличивание) данных перед передачей и хранение эмбедингов в локальной векторной базе в контуре учреждения. Ограничения параметров обуславливаются лимитами длительности видео/аудио и объёма PDF недостаточны для длительных процедур и объёмных историй болезни, что требует сегментации и агрегации эмбедингов. Модель общего назначения может недостаточно дискриминировать медицинскую терминологию и редкие нозологии. Пользовательские инструкции задач частично компенсируют этот недостаток. Может возникнуть проблема интерпретируемости, так как высокоразмерные векторы непрозрачны для клинициста, что требует дополнительных средств объяснимости. Статус публичного превью исключает промышленное использование без валидации.

III. Выводы

Технология обладает характеристиками, перспективными для клинической диагностики: мультимодальность, 3072-мерное пространство с поддержкой MRL, настраиваемые инструкции, встроенный OCR. Основные барьеры – нормативные ограничения, лимиты входных параметров, отсутствие медицинского дообучения и предварительный статус модели. Внедрение рекомендуется через пилотные проекты с обязательным соблюдением принципа «врач в контуре».

Список литературы

1. Google Cloud / Electronic resource. – Mode of access: <https://docs.cloud.google.com/vertex-ai/generative-ai/docs/models/gemini/embedding-2>. – Date of access: 15.03.2026.
2. Закон РБ от 07.05.2021 № 99-3 / Электронный ресурс. – Режим доступа: <https://pravo.by/document/?guid=3871&p0=H12100099>. – Дата доступа: 15.03.2026.

Сальников Даниил Андреевич, ассистент кафедры интеллектуальных информационных технологий БГУИР, d.salnikov@bsuir.by.