

Опыт использования больших языковых моделей для решения задач клиентской аналитики

Артем Летин

Вице-президент, Начальник управления моделирования КИБ и СМБ,
Департамент анализа данных и моделирования, Банк ВТБ (ПАО)

- Согласно опросу СТО*, более половины компаний уже внедрили решения на базе генеративного ИИ в бизнес-процессы
- из них 6% масштабируют эти решения и продают их на внешнем рынке
- Чаще всего решения внедрялись в функциях ИТ, клиентского сервиса, маркетинга и продаж , производства
- Среди популярных кейсов – внедрение технологии в клиентскую поддержку (в чат-боты или в качестве ассистента оператора), персонализация маркетинговых кампаний

Сравнение общей степени внедрения генеративного ИИ в России в 2023 и 2024 году, % компаний среди работающих с генеративным ИИ

Степень внедрения технологий	2023	2024
Проводятся эксперименты по работе с генеративным ИИ в свободном режиме	67%	29%
Определены приоритетные функции и юз-кейсы внедрения генеративного ИИ, решение не внедрено	33%	17%
Решения на основе генеративного ИИ полноценно внедрены хотя бы в одной функции организации	0%	54%

Функции, в которых внедрялись решения на базе генеративного ИИ, % решений

Функции	2023	2024
Производство	9%	36%
HR	11%	37%
Маркетинг и продажи	69%	38%
Клиентский сервис	57%	47%
ИТ	31%	40%
Стратегия	6%	14%
Финансы	3%	13%
Цепочки поставок	3%	14%
Исследования и разработки	49%	28%
Внутренние коммуникации	6%	19%
Среднее число внедрений (среди внедряющих организаций)	2,4	2,9

*Согласно опросу технических директоров (Chief Technology Officers, СТО), проведенному «Яков и Партнёры» летом 2024 г. Опрос включал 100 СТО из топ 300 компаний в каждой из стран в фокусе исследования.

Исследование **techtelligence***: ИИ в клиентской аналитике

- **46%** компаний инвестируют в AI для обслуживания клиентов больше, чем в продажи или маркетинг.
Это связано с сокращением затрат и меньшими рисками в экспериментировании
- более **80%** бизнеса используют виртуальных помощников или «вторых пилотов» в CX, и **75%** из них уже окупают инвестиции
- **50%** контактных центров используют GenAI для генерации ответов клиентам
- **45%** маркетинговых команд применяют GenAI для создания рекламных текстов, постов и сценариев
- **45%** компаний обеспокоены неточностью AI, **39%** — рисками для сохранения и конфиденциальности данных, **31%** — потерей рабочих мест. **61%** респондентов считают, что правительства должны гарантировать право клиентов общаться с живыми людьми

Продукт: Цифровой помощник ВТБ для клиентской аналитики и маркетинга

Задача 1. Работа с продажами:

- Рекомендация продуктов
- Создание резюме встречи/звонка
- Генерация скриптов/Заполнение полей скрипта на основании разговора
- Структурирование текста
- Обратная связь для менеджера по итогам разговора
- Определение тональности беседы
- Формирование контекстных подсказок менеджеру

Задача 2. Работа с «клиентским опытом»:

- Поиск и отбор релевантных отзывов
- Структурирование обращений
- Классификация обращений
- Генерация ответа
- Определение тональности
- Определение болей клиента и направлений развития

Задача 3. Помощь сотруднику КЦ:

- Модель «Суфлер», задача Speech2 text
- формирование контекстных подсказок операторам
- Определение тональности разговора
- Определение срочности вопроса

Задача 4. Вопросно - ответная система:

- Поиск по базе знаний и выдача ответа максимально релевантного заданному контексту

Драйверы: снижение трудозатрат, снижение затрат на персонал, повышение удовлетворенности пользователей, рост NPS (индекс лояльности), рост продаж, сокращение оттока

Первые шаги: Работа с клиентским опытом – повышение эффективности анализа данных соц. медиа

Бизнес требования

Возможность исследования рынка:

- Хотим получить возможность анализировать комментарии для выявления трендов и предпочтений потребителей

Мгновенная обратная связь:

- Хотим анализировать обратную связь от клиентов, автоматически обрабатывать отзывы и предложения пользователей

Цель

- Повысит скорость анализа и обработки запросов
- Улучшение качества обслуживания и продуктов за счет выявления болей клиента и точек роста
- Снизить нагрузку на сотрудников
- Увеличить NPS (индекс лояльности)

Критерии успеха

- Сокращение времени обработки и анализа обращений на 30%
- Положительная ОС от команды клиентского качества
- Увеличение nps на 10пп

Продукт

Трудно говорить об основании, но(!) в программе Гелиос для ИП полис ОСАГО сразу активен, позже подгружаю(в течение одного рабочего дня) платёжное поручение. По банку недоработка: **оплата СБП по бизнес карте у ИП неактивна**, Сбер обещает подключить функционал только в конце года, **ВТБ тоже не позволяет оплачивать СБП по бизнес карте**. Приходится ждать когда поступление денсредств отразится на счёте страховщика, только после этого происходит активация полиса. Очень **хотелось бы как с ООО**. У ИП ведь тоже свой расчётный бизнес счёт.

зоны роста

Проблема



Большой объем текстов

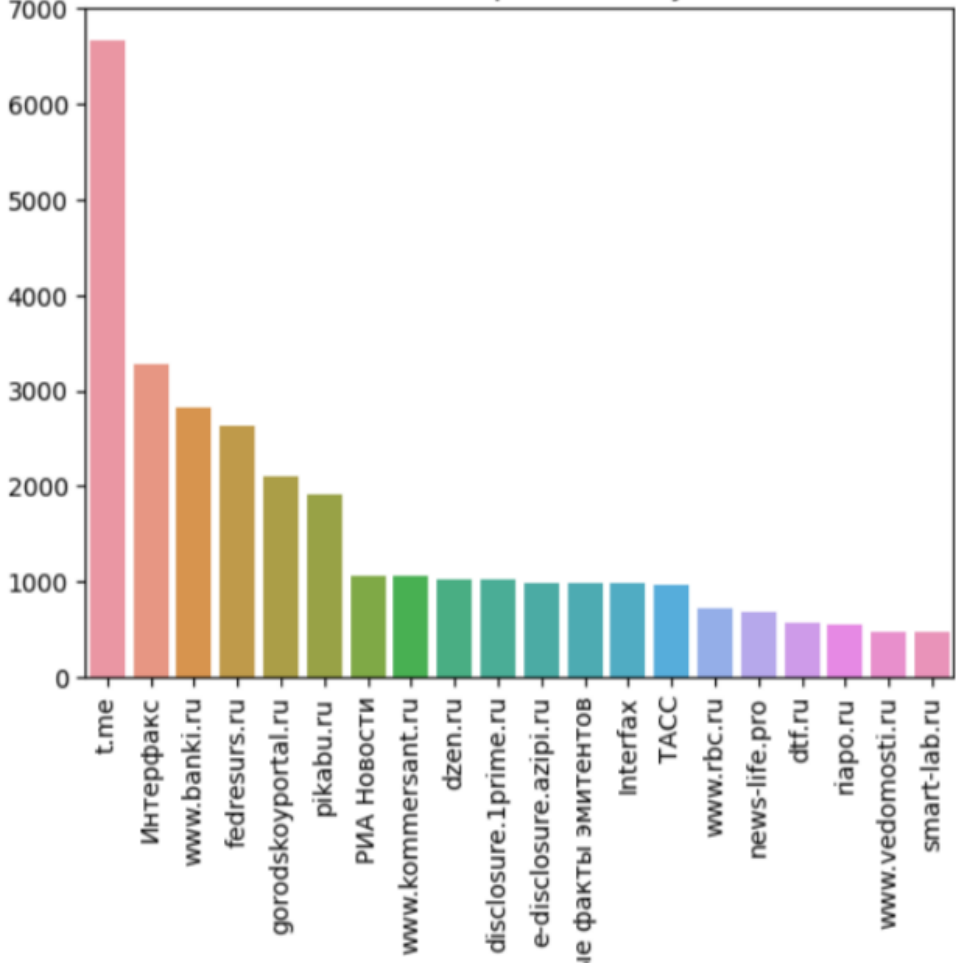


**Неструктурированная
информация**

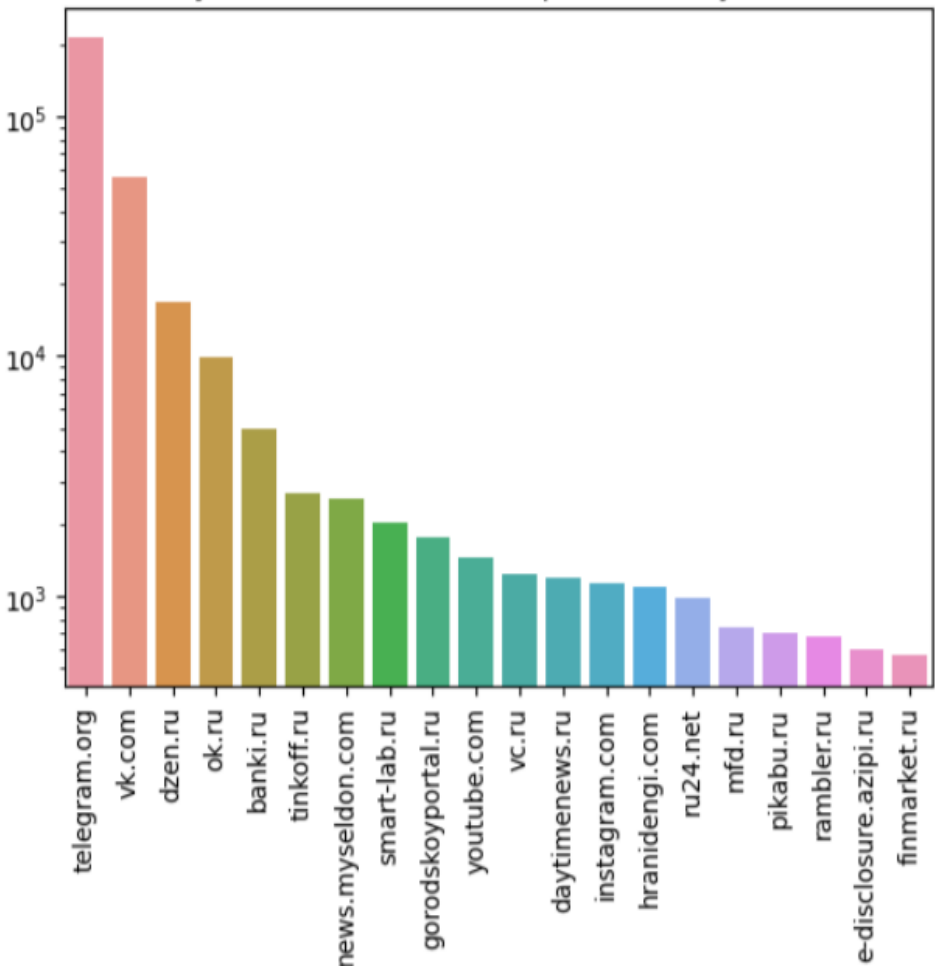


**Поиск нужной
информации глазами**

Скан - 20 наиболее встречаемых суб-источников



Brabdanalytics - 20 наиболее встречаемых суб-источников



Удаление дубликатов и фильтрация спама



Результаты дедупликации

	Август	Сентябрь	Октябрь	Ноябрь
Было	197 586	272 290	358 054	384 583
Стало	92 277	110 200	124 507	136 352



Результаты фильтрации спама

	Август	Сентябрь	Октябрь	Ноябрь
Было	92277	110 200	124 507	136 352
Стало	24 982	26 361	10 553	24 525

Применены:

- нормализация текстов (приведение к стандартной форме, удаление не текстовых символов, пробелов, лемматизация)
- удаление полных дублей нормированных текстов
- векторизация Bert + Tfidf
- алгоритм кластеризации DBSCAN для выявления кластеров с похожими новостями

Применены:

- удаление стоп-слов
- фильтры по стоп-словам, тэгам html
- обучен алгоритм Naive Bayes

Разметка по категориям с использованием LLM

подаем потенциальные категории в промпт и просим LLM выбрать один на основе контекста - комментария

- депозиты для среднего и малого бизнеса (в банке есть депозиты: "Авансовый", "Гибкий", "Залоговый", "Овернайт", "Универсальный", "Фиксированный", Копилка для бизнеса, вексель)
- банковские гарантии и аккредитивы среднего и малого бизнеса (например инкассо, Гарантии участникам закупок 44-ФЗ, 223-ФЗ, 615-ПП)
- кредитование малого и среднего бизнеса (в том числе Программа №2221, Программа кредитования «Урожайный бизнес», Программа №1575, Программа №2116, Кредит на Госзаказ, Кредит на залоговое имущество, Программа №445, Кредитование жилищного строительства, Кредитование недропользователей, Льготное кредитование сельхозпроизводителей, Оборотный кредит, Овердрафт, Финансирование капитальных затрат)
- регистрация бизнеса
- бизнес карты для среднего и малого бизнеса (в том числе Таможенная карта, Универсальная карта для бизнеса)
- дистанционное банковское обслуживание среднего и малого бизнеса (в том числе Интеграционный Банк-Клиент, Интернет-банк "Бизнес Платформа ВТБ", Мобильный банк «Бизнес Платформа ВТБ», 1С: ДиректБанк)

Отдельная псевдокатегория «банковское обслуживание и продукты для физических лиц» чтобы модель более четко понимала, что мы выделяем продукты для юрлиц

Карта Н100, категоризация
23 000 сообщений за 10 часов



Qwen 2.5-
32B-Instruct

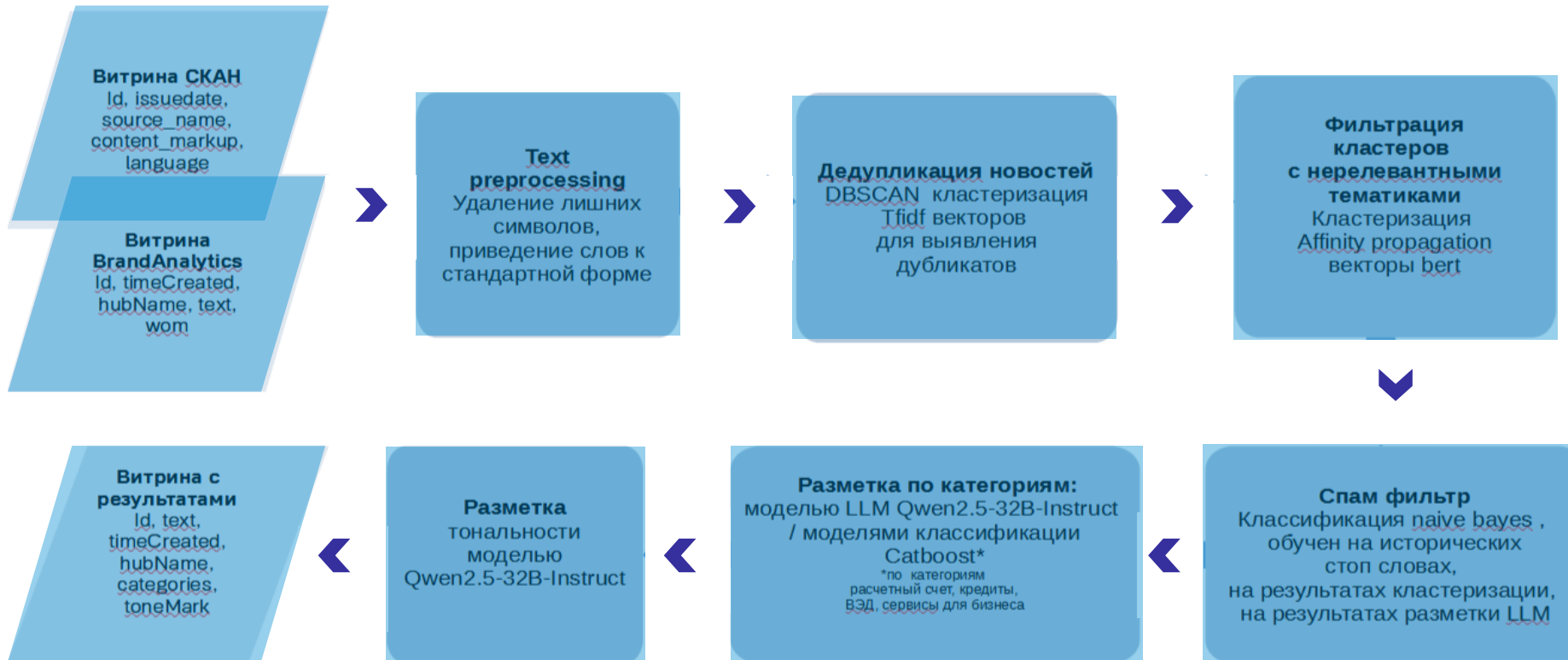
Длина
контекста:
131,072 tokens
и генерация
8192 tokens

Техника Few shot:

брали релевантное
сообщение и показывали как
пример

	f1_score	recall	precision
Расчетный счет	0.89	0.86	0.92
Кредиты	0.87	0.85	0.89
ВЭД	0.86	0.75	1.0
Сервисы для бизнеса	0.77	0.71	0.83

Схема процесса разработки решения



Примеры категорий

level_1	level_2
проблема банка	отказ от работы с ВТБ/негатив
	стабильность банка
	надежность банка
проблема персонала	невежливое обращение
	незаинтересованность в решении вопроса/Вопрос не решен
	отказ в обслуживании
	удобство обслуживания
	неверная/неполная консультация
проблема продукта	сроки
	ставка
	условия
	порядок оформления
	удобство пользования
	несогласие с комиссией
	некорректное списание, начисление
	услуга не получена
	не проинформирован об изменении условий
	подключен без согласия клиента
прочее	

Образ результата: что можно выводить в BI

id	text	category	tone_vtb	tone_product	summary	level1	level2	growth
14160259	Позвонил в контактный центр, попросил открыть р/с для моего ИП, прошел банковскую проверку по телефону (ожидая несколько минут), после чего оператор мне сообщил что проверка пройдена положительно, мне необходимо лично прийти в ближайшее отделение банка и назначил сразу время, я поехал туда там мне сразу сообщили что у меня отказ в открытии р/с и попробовали завести заявку повторно в ручную, на которую также был отказ. Прошу посодействовать в открытии счета для моего ИП, при учете что у меня ИП открыто 4 года назад, есть 4 действующих счета в других банках (сбер, втб, точка, уралсиб), и у налоговой ко мне нет вопросов (я им звонил - спрашивал есть ли ко мне вопросы, тк я не могу открыть счет, они ответили что вопросов ко мне нет и задолженностей по налогам тоже!). также прошу прояснить мне на каком основании мне отказывают в открытии счета	Расчетный счет	-1	-1	ИП столкнулся с отказом в открытии расчетного счета в банке несмотря на наличие нескольких действующих счетов в других банках и отсутствие задолженностей перед налоговой	проблема продукта	порядок оформления	Улучшение процесса информирования пользователей
13744683	Банк установил для организаций повышенную комиссию за переводы физлицам. Компания посчитала это заградительным тарифом, который взимают без объективных причин. В 2021 году компания «XXX» открыла счет в ВТБ. По правилам комплексного банковского обслуживания банк взимает комиссию 10%, если организация переводит на счет физлиц в другом банке больше 5 млн руб. В 2023 году «XXX» перевел 10 млн руб. по договору займа XXXX. Банк списал за эту операцию 10% комиссии, то есть 1 млн руб. Компания посчитала, что ВТБ списал деньги необоснованно, и сначала обратилась в банк. Кредитная организация частично вернула комиссию и выплатила организации 200 000 руб. После этого компания обратилась в суд (дело №XXXX/XXXX). «XXXX» требовал взыскать с банка 800 000 руб. неосновательного обогащения и проценты за пользование чужими деньгами. Суды отказали организации в удовлетворении требований. Они указали, что банк списал комиссию в рамках договора о комплексном банковском обслуживании. Компания знала условия обслуживания и добровольно подписала соглашение с кредитной организацией. В кассационной инстанции истец дополнительно заявил, что у такой повышенной комиссии нет экономического обоснования. Но этот довод суд тоже отклонил, потому что компания не заявляла его в других инстанциях. В Верховном суде «XXXX» доказывала, что повышенная комиссия относится к заградительному тарифу, который экономически не оправдан. Экономколлегия поддержала доводы компании. Вводя повышенные тарифы на переводы физлицам, банк заведомо недобросовестно пользуется своими правами и подрывает ожидания клиентов. ВС установил, что комиссия «имеет очевидные признаки заградительного тарифа, препятствующего совершению законной банковской операции клиента». Комиссия не учитывает реальность и экономическую обоснованность. На это нужно обратить судам при новом рассмотрении дела.	Расчетный счет	-1	-1	Компания «Имотэк» оспаривает в суде повышенную комиссию за переводы физическим лицам с расчетного счета, считая ее необоснованной и заградительным тарифом	проблема продукта	несогласие с комиссией	Уменьшить комиссии и повысить прозрачность

Следующие шаги: Интерфейс взаимодействия с пользователем

Привет, я цифровой помощник ВТБ. Я могу предоставить анализ по тематикам:

1. Какая отношение к рассматриваемому банку на основании отзывов (с эмоциональной окраской);
2. Какое отношение к продукту – мнение (с эмоциональной окраской);;
3. Какое условия продукта не нравятся – ставки, сроки и т.д.;

Какие зоны роста по продукту ты видишь?



Уточните, пожалуйста, период за который требуется аналитика

За последний месяц

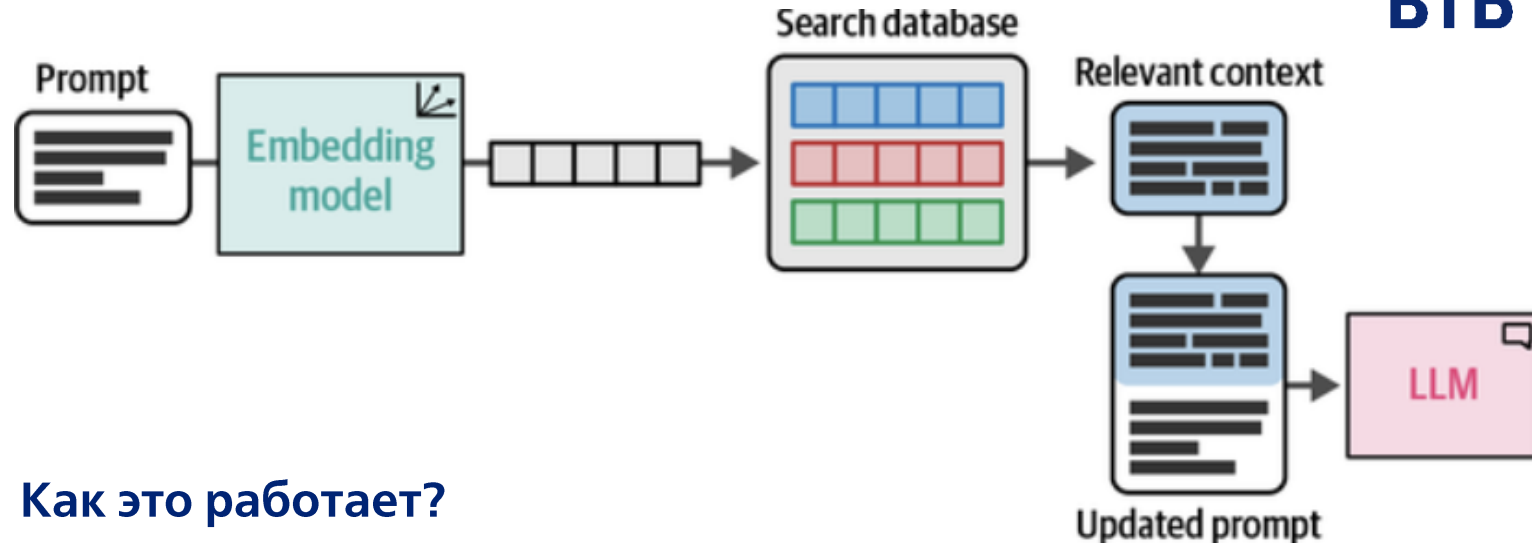


Можно изменить условия продукта – снизить требования к заемщику, изменить срок рассмотрения заявок, сделать плату за обслуживание дешевле

Retrieval Augmented Generation (RAG)

– это технология, сочетающая в себе поиск релевантной информации в существующих хранилищах данных и генерацию текста с помощью языковых моделей, для создания более точных информативных ответов.

В качестве данных (на схеме search database) может выступать любое хранилище информации, по которому можно производить поиск. Например, база данных службы поддержки с парами вопрос-ответ, документация компании, база с текстами из энциклопедии, база с персональными данными пользователя.



Как это работает?

1. Для каждой категории проблем создается сводное описание, которое меняется во времени
2. Пользователь задает вопрос, например, «Какие зоны роста по продукту?»
3. Далее вопрос переводится в вектор
4. В векторной базе ищутся по семантической (смысловой) близости похожие вектора
5. Далее на основе найденных данных автоматически формируется вопрос к LLM, например, «Ответь на вопрос на основе приложенных данных: какие зоны роста по продукту?»
6. Далее LLM генерирует пользователю ответ.

Фактически, по такой схеме можно «общаться» с любой корпоративной базой знаний.

Основные проблемы

- _____
- _____
- _____

Зоны роста

- _____
- _____
- _____

Плюсы относительно конкурентов

- _____
- _____
- _____

LLM не заменяет весь пайплайн, но ускоряет процесс разработки

LLM расширяет возможности автоматизации: использование модели делает возможность создания более интерактивного взаимодействия с пользователем и за счет уточняющих вопросов. В перспективе можно интегрировать дополнительные данные о ЮЛ (отзывы о продукте, банке) для создания персонализированного предложения

Для интеграции **LLM** в существующую экосистему нужны существенные доработки, которые по сути означают создание цифрового продукта

LLM не должен заменять, а должен дополнять существующие решения