



АФИИНА — универсальный стандарт оценки финансового эффекта от ИИ

Открытая база знаний для компаний всех отраслей. Методология, метрики, формулы расчёта и типовые кейсы оценки эффективности ИИ-инициатив.

Содержание

АТЛАС ФИНЭФФЕКТОВ ИИ: НОВЫЕ АКТИВЫ

О ПРОЕКТЕ

О проекте АФИИНА	6
Единый язык оценки ИИ	6
Команда проекта	7
Команда управления и методологии проекта	7
Рабочая группа проекта	7

МЕТОДОЛОГИЯ

01	Методология создания	8
	Как создавалась АФИИНА?	8
	Трёхуровневая модель оценки ИИ	8
	АФИИНА I — Финансовый уровень	8
	АФИИНА II — Стратегический уровень	9
	АФИИНА III — Системный уровень	9
	Сводная таблица параметров уровней	10
02	Определение ИИ	11
	Общий пример для всех отраслей	11
	1. Финансовая индустрия	11
	2. Ритейл и e-commerce	11
	3. Транспорт и логистика	12
	4. АПК, энергетика и промышленность	12
	5. ИКТ	12
03	Процесс оценки	13
	АФИИНА I — Краткосрочный уровень (Финансовый)	14
	АФИИНА II — Среднесрочный уровень (Стратегический)	14
	АФИИНА III — Долгосрочный уровень (Системный)	15
	Процесс оценки: этапы	16

04	Принципы	17
	Общие принципы	17
	Объект оценки: AS IS/TO BE	21
	Инкремент	23
	Чистый vs Валовый эффект	25
05	Метрики	27
	Типы финансовых эффектов	27
	Набор метрик для оценки финансового эффекта в зависимости от целей, задач	37
	АФИИНА I	37
	АФИИНА II	38
	II.I. Финансовая индустрия	40
	II.II. Ритейл и e-com	41
	II.III. Транспорт и логистика	42
	II.IV. ИКТ	43
	II.V. АПК, энергетика и промышленность	44
	АФИИНА III	46
	III.I. Финансовая индустрия	47
	III.II. Ритейл и e-com	48
	III.III. Транспорт и логистика	49
	III.IV. ИКТ	49
	III.V. АПК, энергетика и промышленность	50
	R&D	50
	Бенчмаркинг	53
06	Корректность расчета	61
	ТОП-ошибок при расчете эффекта	61
	Двойная оценка фин.эффекта	65
	Классы потенциальных задвоений	65
	Методика оценки двойного эффекта	67
	Финансовая индустрия	68
	Ритейл и e-com	68
	Транспорт и логистика	69
	ИКТ	69
	АПК, промышленность и энергетика	70

07	Методы	71
	АФИИНА I	71
	I.I. Базовый A/B-тест	71
	I.II. Альтернативное моделирование	76
	I.III. Отраслевая специфика	77
	АФИИНА II	78
	II.I. Усиленный A/B тест	78
	II.II. Квази-экспериментальные методы	81
	II.III. Альтернативное моделирование	83
	II.IV. Отраслевая специфика	83
	АФИИНА III	84
	III.I. Альтернативное моделирование на исторических и рыночных данных	84

КЕЙСЫ

01	Финансовая отрасль	89
	01 Кампании продаж и модели приоритизации клиентов на основе ИИ	89
	02 Чат-бот клиентского обслуживания	90
	03 Модель определения дохода клиента в целях расчета Показателя Долговой Нагрузки	91
	04 Прогнозирование метрик регулярной отчетности с помощью AutoML	92
	05 AI-агент по формированию one-page-слайда по презентации	93
	06 Скоринг для исходящих коммуникаций с клиентами банка в Контактном центре	94
	07 Создание AI-аватара для сотрудников поддержки	95
	08 Чат и голосовой бот клиентского обслуживания	96
	09 Модель привлечения BARCELONA (продукт Премиум)	97
	10 Аналитический сервис по автоматизации процесса оценки и переоценки отдельных видов профильных залогов юридических лиц	97
	11 Время на поиск ответов по бизнес Википедии и продуктовому каталогу для сотрудников колл центра (КЦ)	99
02	Ритейл и e-com	100
03	ИКТ	101
	01 Видеоаналитика с применением ИИ при обслуживании инфраструктурных объектов	101
	02 AI Engineer для автоматизации инцидент-менеджмента сети	102

04	Промышленность, энергетика и АПК	104
	ИИ-агент подбора материалов из запасов компании	104
05	Транспорт и логистика	105
	Единая платформа коммуникаций и чат-бот клиентской поддержки	105

О проекте АФИИНА

Единый язык оценки ИИ

АФИИНА (Атлас финэфектов ИИ: новые активы) — открытый wiki-портал, созданный для формирования универсального стандарта оценки финансового эффекта от внедрения технологий искусственного интеллекта.

Проект основан на **Методологии оценки финансовой эффективности от ИИ** — совместной разработке Альянса в сфере ИИ и Ассоциации ФинТех, подписанной ведущими финансовыми организациями на конференции AI Journey 2025.

Методология помогает объективно оценивать, обосновывать и масштабировать ИИ-инициативы. В ней описаны подходы к расчёту выгоды от коммерческих проектов, R&D-инициатив и решений без прямой мгновенной прибыли.

Первоначально созданная для финансовой отрасли, методология расширяется для применения **во всех секторах экономики**.

«Общие стандарты эффективности и единые метрики позволят всем игрокам говорить на одном языке при внедрении ИИ, что повысит прозрачность инвестиций и доверие к технологиям.»

Александр Ведяхин

Первый зам. Председателя Правления Сбербанка, председатель наблюдательного совета Альянса в сфере ИИ

ПОДПИСАНТЫ МЕТОДОЛОГИИ

Сбер, Т-Банк, Альфа-Банк, Россельхозбанк, Газпромбанк, Московская биржа, ВСК, Дом.РФ, МКБ, ВЭБ РФ, Капитал Лайф Страхование

Команда проекта



Валерия Воробьева

Генеральный директор Альянса ИИ

«Искусственный интеллект стал полноценным бизнес-активом, но чтобы масштабировать его внедрение, рынку нужны единые правила оценки эффекта. Прорыв АФИИНЫ — в создании универсального стандарта сразу для пяти групп отраслей: от е-ком и ритейла до промышленности и АПК. Мы делаем ее не статичным документом, а динамической методологией, которая будет расти и меняться вместе с рынком, вобрав в себя лучший опыт отраслевых экспертов»

Команда управления и методологии проекта

// АЛЬЯНС В СФЕРЕ ИИ



Никита Худов

Заместитель генерального директора

РАНЕЕ

Исполнительный директор и глава
Центра AI-трансформации Сбера

// АЛЬЯНС В СФЕРЕ ИИ



Анастасия Пименова

Руководитель проектов

РАНЕЕ

Руководитель проектов департамента
развития ИИ Сбера

Рабочая группа проекта

// АЭРОФЛОТ



Светлана Потапова

Директор Департамента по развитию
искусственного интеллекта

// АЭРОФЛОТ



Екатерина Попова

Начальник отдела внедрения решений
и проектного управления

// БИЛАЙН



Татьяна Крылова

Директор по трансформации ИИ

// КУПЕР



Элен Теванян

Head of ML

// НОРНИКЕЛЬ



Наталья Артамонова

Директор по внедрению генеративного
ИИ

// НОРНИКЕЛЬ



СМ Сергей Мамзин

Руководитель направления по
внедрению AI агентов

// РОСТЕЛЕКОМ



Анна Малиновская

Директор направления, департамент по
внедрению ИИ продуктов и технологий

Методология создания

Методология создания

Как создавалась АФИИНА?

АФИИНА (Атлас Финансовых Эффектов ИИ: Новые Активы) — открытый wiki-портал и методологический стандарт оценки финансового эффекта от внедрения ИИ. АФИИНА — проект Альянса в сфере ИИ.

В основе АФИИНЫ лежит Методология оценки финансовой эффективности от ИИ — документ, разработанный при участии ведущих финансовых организаций страны и представленный на конференции AI Journey 2025. Среди подписантов — Сбер, Т-Банк, Альфа-Банк, Газпромбанк, Московская биржа и другие.

Методология изначально создавалась для финансового сектора — отрасли с наиболее зрелой культурой оценки инвестиций и наибольшим числом ИИ-внедрений. При этом она с самого начала проектировалась как универсальный стандарт, применимый во всех отраслях: ритейле, логистике, промышленности, телекоме, девелопменте и других.

Структура АФИИНЫ состоит из двух частей:

- **Базовая (унифицированная) часть** — принципы, метрики и методы оценки, общие для всех отраслей.
- **Отраслевая часть** — специфика, особенности расчётов и кейсы для конкретных секторов.

Основной инструмент структурирования методологии — трёхуровневый фрейм оценки. Он определяет, с какой точки зрения и для каких целей оценивается эффект от ИИ.

Трёхуровневая модель оценки ИИ

Оценка эффекта от ИИ зависит от уровня зрелости компании и её целей. Фрейм выделяет три горизонта оценки — от финансового, на котором сегодня работает большинство компаний, до системного, который применяется пока редко.

Уровни преемственны: каждый следующий включает предыдущий. При переходе на уровень выше компания сохраняет уже используемые метрики, но применяет их в большем масштабе. От уровня зависят цель оценки, горизонт планирования, состав участников и инструментарий.

АФИИНА I — Финансовый уровень

Ключевой вопрос: сколько денег принесла или сэкономила конкретная ИИ-инициатива?

Базовый уровень, с которого начинают большинство компаний. Оценивается один ИИ-продукт или процесс и его прямая отдача в пределах года. Задача — выразить эффект в деньгах и отразить его в отчёте о прибылях и убытках (PnL). Уровень нужен для внешней отчётности и обоснования инвестиций перед руководством.

- **Горизонт:** до 1 года

- **Объект оценки:** один ИИ-продукт или процесс
- **Метрики:** PnL отчётного года, PnL на 12 месяцев вперёд
- **Методы:** базовый А/Б тест, альтернативное моделирование
- **Роли:** бизнес-юнит

АФИИНА II — Стратегический уровень

Ключевой вопрос: как ИИ повышает конкурентоспособность компании?

Уровень для компаний, которые оценивают не отдельные проекты, а весь портфель ИИ-инициатив. Горизонт расширяется до нескольких лет. Помимо прямой прибыли учитываются изменение клиентского опыта, рост долгосрочной ценности клиента (LTV) и ускорение операционных циклов. Это позволяет принимать обоснованные инвестиционные решения и управлять портфелем технологий.

- **Горизонт:** 2–5 лет
- **Объект оценки:** портфель ИИ-инициатив компании
- **Метрики:** NPV, CLTV, ROI с драйверной моделью, Time-to-XXX, CSI/NPS
- **Методы:** продвинутый А/Б тест, вечный А/Б тест, квази-эксперименты, альтернативное моделирование
- **Роли:** бизнес-юнит + финансовая валидация + техническая валидация

АФИИНА III — Системный уровень

Ключевой вопрос: как массовое внедрение ИИ меняет отрасль и экономику?

Уровень оценки макроэффекта в ситуации, когда ИИ внедряют все участники рынка одновременно; на практике применяется пока редко. Использует метрики двух предыдущих уровней и добавляет отраслевой и макроэкономический срез: как меняются отраслевые риски, требования к капиталу и устойчивость системы, в которой работают схожие алгоритмы. Уровень ориентирован на регуляторов и крупные компании.

- **Горизонт:** 5+ лет
- **Объект оценки:** отрасль или экономика в целом
- **Метрики:** резервы, капитал, системная устойчивость и риски
- **Методы:** альтернативное моделирование на исторических и рыночных данных
- **Роли:** бизнес-юнит + финансовая валидация + аудит + ИИ-стратеги

Сводная таблица параметров уровней

ПАРАМЕТР	АФИИНА I: ФИНАНСОВЫЙ	АФИИНА II: СТРАТЕГИЧЕСКИЙ	АФИИНА III: СИСТЕМНЫЙ
Ключевой вопрос	Окупилась ли инициатива?	Как ИИ усиливает компанию?	Как ИИ меняет отрасль?
Горизонт	До 1 года	2–5 лет	5+ лет
Объект	Один продукт или процесс	Портфель инициатив компании	Отрасль или экономика
Метрики	PnL	NPV, CLTV, ROI, NPS, Time-to-XXX	Резервы, капитал, риски
Методы	Базовый А/Б тест	Продвинутый А/Б, квази-эксперименты	Альтернативное моделирование
Роли	Бизнес-юнит	+ Фин. и Тех. валидация	+ Аудит, ИИ-стратегия

Определение ИИ

Что такое ИИ?

ИИ/МЛ – это алгоритмы на базе обучающейся математической модели, которые позволяют выполнять задачи, требующие человеческого интеллекта, на основе анализа данных

Эффект от ИИ/МЛ возникает только при внедрении модели в бизнес-процесс для решений поставленных задач.

Ниже приведены примеры применения ИИ. Первый пример — общий для всех отраслей, остальные показывают специфику конкретных секторов экономики.

Общий пример для всех отраслей

Клиентский сервис — ИИ-обработка обращений в контактном центре

ИИ-ассистент в контактном центре (КЦ) принимает обращения клиентов, отвечает на типовые вопросы и маршрутизирует сложные запросы на профильных специалистов. Компания обрабатывает больше обращений тем же штатом, сокращает время ответа и снижает нагрузку на операторов. Такой сценарий применим в любой отрасли — от банка до промышленного предприятия.

1. Финансовая индустрия

ИИ-скоринг в банке

Модель анализирует сотни параметров заёмщика — от истории транзакций до поведения в мобильном приложении — и за несколько секунд выдаёт решение о выдаче кредита. Банккратно ускоряет рассмотрение заявок, снижает финансовые риски и освобождает кредитных аналитиков от рутинных проверок.

2. Ритейл и e-com

ИИ-персонализация в онлайн-магазине

Алгоритм непрерывно изучает историю покупок и просмотров пользователя, чтобы в наиболее удачный момент предложить релевантный товар или индивидуальную скидку. Покупатели чаще возвращаются в магазин и тратят больше, а маркетинговые бюджеты расходуются эффективнее.

3. Транспорт и логистика

ИИ-маршрутизация в доставке

Алгоритм отслеживает геолокацию курьеров, пробки и поток новых заказов, формируя для водителей оптимальные маршруты в режиме реального времени. Компания экономит на топливе и успевает развезти больше заказов тем же штатом курьеров.

4. АПК, энергетика и промышленность

ИИ-предиктивное обслуживание на производстве

Датчики на оборудовании передают данные в ИИ-модель, которая распознаёт скрытые аномалии и предсказывает поломку за несколько дней до её возникновения. Завод отказывается от ремонта строго по календарю и обслуживает технику только тогда, когда это действительно нужно, избегая внезапных остановок конвейера.

5. ИКТ

ИИ-управление сетью

Интеллектуальная система отслеживает перепады трафика и автоматически перенаправляет мощности на перегруженные участки сети до того, как связь начнёт обрываться. Оператор сохраняет высокое качество связи в часы пик, минимизирует аварии и снижает затраты на выезды ремонтных бригад.

Процесс оценки

Процесс: зачем и кто оценивает фин. эффект в рамках ИИ трансформации?

Зачем оценивать: цели оценки финансового эффекта

Оценка финансового эффекта от ИИ решает как внутренние задачи компании, так и внешние — перед рынком и акционерами.

Внутри компании:

- **Обоснование портфеля ИИ-инициатив** — оценка ROI и приоритизация инициатив. Для классического МЛ оценка строится на ROI и прибыли до налога ближайшего года или трёх лет; для генеративного ИИ (ГенИИ) — на более долгосрочных метриках, например CLTV.
- Проверка соответствия инициатив стратегическим целям компании.
- **Приоритизация ресурсов для внедрения ИИ** — инфраструктуры и людей.
- **Управление бизнес-рисками** — например, в части продаж и производительности.

Для внешнего рынка:

- Усиление репутации.
- Увеличение стоимости бренда.
- Привлечение и удержание акционеров.

Кто оценивает: роли и ответственность

Состав участников процесса оценки определяется **уровнем зрелости компании** в части ИИ. Методология предусматривает три уровня погружения — каждый задаёт обязательный и опциональный набор ролей.

Описание ролей

- **Бизнес-юнит** — владелец ИИ-инициативы внутри продукта или процесса. Формулирует гипотезы, рассчитывает финансовый эффект, отвечает за внедрение модели и достижение результата.
- **Финансовая валидация** — независимый финансовый блок. Проверяет корректность расчёта эффекта, допущений и его отражение в PnL и бизнес-плане.
- **Техническая валидация** — проверяет корректность ML-решения, дизайна экспериментов (A/B, квази-эксперименты), качество данных и моделей.
- **ИИ-стратегия** — проверяет соответствие инициативы (модели) стратегическим целям компании; участвует в приоритизации портфеля.
- **Аудит** — внутренняя, полностью независимая функция, которая оценивает работу всех остальных ролей: качество их работы, нарушения и отсутствие предвзятости. Аудит проверяет всех, но не отвечает за выполнение шагов процесса.

Чем выше уровень зрелости, тем шире круг участников и тем более полным и верифицированным становится результат оценки.

АФИИНА I — Краткосрочный уровень (Финансовый)

Цель данного уровня: внешняя отчётность и подсчёт окупаемости ИИ-инициатив.

Этот уровень подходит компаниям, только начинающим системно оценивать финансовый эффект от ИИ, а также для ситуаций, когда требуется быстро обосновать инвестиции перед стейкхолдерами. Оценку проводит бизнес-юнит самостоятельно — внутри продукта или процесса.

Роли и ответственные:

роль	участие
Бизнес-юнит (внутри продукта/процесса)	+
Финансовая валидация	+ —
Техническая валидация	+ —
ИИ-стратегия	+ —
Аудит	+ —

+ - Обязательные роли

+/- - Опциональные роли

- **Бизнес-юнит (основная роль).** Самостоятельно формулирует гипотезу, рассчитывает окупаемость инициативы в PnL и готовит данные для внешней отчётности.

Остальные роли (финансовая и техническая валидация, ИИ-стратегия, аудит) на этом уровне малоактивны и привлекаются точно по запросу.

АФИИНА II — Среднесрочный уровень (Стратегический)

Цель данного уровня: повышение конкурентоспособности компании.

Этот уровень применяется компаниями со сформированными ИИ-командами, которые управляют портфелем инициатив и заинтересованы в верификации результатов независимым финансовым блоком.

роль	участие
Бизнес-юнит	+
Финансовая валидация	+
Техническая валидация	+
ИИ-стратегия	+
Аудит	+

+ - Обязательные роли

+/- - Опциональные роли

Описание ролей на текущем уровне:

- **Бизнес-юнит.** Оценивает портфель инициатив, строит драйверные модели, отвечает за расчёт эффекта.
- **Финансовая валидация.** Независимо валидирует эффект, допущения и его отражение в PnL и бизнес-плане.
- **Техническая валидация.** Проверяет качество моделей и корректность дизайна экспериментов (A/B, квази-эксперименты).
- **ИИ-стратегия.** Проверяет соответствие инициатив стратегическим целям компании, приоритизирует портфель.
- **Аудит.** Независимо контролирует качество работы всех функций, нарушения и отсутствие предвзятости.

АФИИНА III — Долгосрочный уровень (Системный)

Цель данного уровня: повышение конкурентоспособности и эффективности отрасли / экономики в целом.

Этот уровень характерен для компаний с высокой степенью ИИ-зрелости, рассматривающих ИИ как источник долгосрочных стратегических преимуществ. На данном уровне требуется полный состав участников: оценка затрагивает фундаментальные показатели и требует независимой верификации с нескольких сторон.

роль	участие
Бизнес-юнит	+
Финансовая валидация	+
Техническая валидация	+
ИИ-стратегия	+
Аудит	+

+ - Обязательные роли

+/- - Опциональные роли

- Состав ролей тот же, что на АФИИНА II; меняется специфика задач — оценка ведётся на уровне отрасли/системы.
- **Бизнес-юнит.** Оценивает вклад ИИ на уровне бизнес-направления или отрасли.
 - **Финансовая валидация.** Валидирует эффект на исторических и рыночных данных, фундаментальные показатели (резервы, капитал).
 - **Техническая валидация.** Проверяет системные модели и контрфактические сценарии.
 - **ИИ-стратегия.** Оценивает соответствие ИИ-стратегии отраслевым и системным целям.
 - **Аудит.** Полный независимый надзор за достоверностью системного эффекта, нарушениями и отсутствием предвзятости.

Процесс оценки: этапы

Процесс показан на примере уровня АФИИНА II, где активны все роли. На уровне АФИИНА I этот же процесс ведёт практически один бизнес-юнит, а на уровне АФИИНА III состав ролей сохраняется, но задачи решаются на отраслевом/системном уровне.

Оценка проходит пять последовательных этапов (по жизненному циклу расчёта эффекта). На каждом этапе закреплена зона ответственности ролей — см. RACI-матрицу ниже.

- **1. Формулировка бизнес-гипотез.** Бизнес-юнит формулирует гипотезу эффекта и закладывает консервативный сценарий (минимальные доходы, максимальные расходы).
- **2. Первичная оценка финансового эффекта.** Начальная проверка предположений и прогноз влияния ИИ на бизнес-процессы; финансовая валидация подтверждает корректность подхода.
- **3. Проведение пилота.** Проверка гипотез в реальных условиях (A/B-тест и др.); техническая валидация отвечает за корректность дизайна эксперимента.
- **4. Оценка эффекта.** Точный анализ результатов пилота и расчёт вклада ИИ в PnL; финансовая валидация подтверждает итог.
- **5. Мониторинг эффекта.** Постоянный контроль драйверов, выявление деградации модели и переоценка при необходимости.

End-to-end схема процесса



Принципы

Основные принципы оценки

Принципы, изложенные в данной главе, являются универсальными и обязательными для компаний любого уровня зрелости — вне зависимости от того, на каком этапе ИИ-трансформации находится организация. Они одинаково применимы как при краткосрочной финансовой оценке в рамках первого уровня фрейма (АФИИНА I), так и при долгосрочном системном анализе влияния ИИ на бизнес в рамках третьего уровня (АФИИНА III). Соблюдение этих принципов — необходимое условие любой качественной оценки финансового эффекта от ИИ: без их выполнения результаты расчётов теряют объективность, сопоставимость и управленческую ценность.

Общие принципы

Важно понимать, что финансовая оценка – это не набор универсальных математических законов и операций. Финансовая оценка – это обоснованное мнение, подкреплённое расчётом. По определению, мнение подразумевает субъективность, которую необходимо минимизировать для принятия взвешенных управленческих решений.

Для этой цели разработан набор из 7 принципов финансовой оценки, применимых к ИИ-инициативам в компаниях любой отрасли.

01 Принцип максимальной объективности

В процессе финансовой оценки необходимо избегать экспертных суждений, не подкреплённых расчётами. Оценка должна опираться на объективные данные и статистику. Иными словами, любой, кто располагает той же статистикой, должен иметь возможность самостоятельно воспроизвести расчёт и выявить реальное влияние ИИ-инициативы на бизнес-показатели.

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

эффект кредитно-скоринговой модели подтверждается статистикой одобрений, дефолтов и марже по каждому сегменту — без ссылок на «экспертное мнение» аналитика

Ритейл и e-com

эффект рекомендательной системы подтверждается данными A/B теста по конверсии и среднему чеку, а не оценкой команды по принципу «мы видели рост»

Транспорт и логистика

снижение расходов на топливо от ИИ-маршрутизации рассчитывается на основе телематических данных по каждому маршруту, а не на усреднённом экспертном нормативе

АПК, энергетика и промышленность

эффект от ИИ-предиктивного обслуживания оборудования подтверждается статистикой простоев, общими затратами компании на аварийные ремонты на целевом контуре, а также дельтой недополученной выручки от простоев до и после внедрения

ИКТ

влияние модели прогнозирования оттока клиентов на их удержание оценивается путем сравнения фактического уровня оттока (churn) в двух группах - одной, где используется модель, и другой, где она не применяется. Это позволяет точно измерить, насколько эффективно использование ИИ-модели способствует снижению числа уходящих клиентов и улучшению их удержания.

02 Принцип измеримости

Выгоды и затраты по ИИ-инициативе должны быть измеримы. Все параметры расчёта фиксируются в качестве **драйверов** — как финансовых (операционный доход, OPEX, свободный денежный поток), так и нефинансовых (количество продаж, отклик, текучесть персонала). В любой момент жизненного цикла инициативы необходимо иметь возможность получить актуальное значение каждого драйвера.

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

драйверы эффекта антифрод-модели — объём выявленного мошенничества (руб.), доля ложных срабатываний, операционные расходы на расследования

Ритейл и e-com

драйверы персонализации выдачи — конверсия в покупку, средний чек, доля повторных покупок — мониторятся еженедельно

Транспорт и логистика

драйверы ИИ-планирования загрузки — коэффициент заполненности транспортного средства, километраж вхолостую, затраты на топливо на км

АПК, энергетика и промышленность

драйверы ИИ-мониторинга качества продукции — доля брака на линии, стоимость одного дефекта, удельные нормы расхода сырья и материалов на единицу годной продукции

ИКТ

драйверы модели оттока — число клиентов в зоне риска, стоимость удержания одного клиента, ARPU и средний ежемесячный доход, который компания получает от удержанного клиента

03 Принцип консерватизма

Денежные потоки обладают свойством неопределённости. Необходимо выбрать единый подход к работе с неопределённостью, чтобы финансовые оценки можно было сравнивать между собой. Чтобы минимизировать риски реализации негативных сценариев и тем самым повысить доверие к оценке, принцип консерватизма предлагает на этапе планирования для всех ИИ-инициатив закладывать максимально негативный сценарий: **доходы оцениваются на минимально возможном уровне, расходы — на максимально возможном.**

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

при оценке кредитной модели прогноз прироста одобрений берётся по нижней границе доверительного интервала А/Б теста, а стоимость дефолтов — по верхней

Ритейл и e-com

в оценке ИИ-ценообразования закладывается минимальный прирост маржи и максимальная скидочная нагрузка

Транспорт и логистика

экономия топлива от умной маршрутизации планируется исходя из минимального значения, зафиксированного в пилотном регионе, — без экстраполяции на «лучший» регион

АПК, энергетика и промышленность

эффект от снижения брака оценивается по минимально подтверждённому снижению дефектности, с учётом сезонных колебаний качества сырья

ИКТ

прогноз удержания от модели оттока строится по пессимистичному сценарию: средний доход от одного пользователя (ARPU) у возвращённых клиентов оценивается с учётом того, что они могут не сразу вернуться к своим прежним уровням потребления услуг или могут использовать менее прибыльные тарифы.

04 Принцип связи с инициативой

В качестве выгод и затрат рассматриваются **инкрементальные (дополнительные) потоки**, связанные именно с оцениваемой ИИ-инициативой. Если инициатива является не единственным фактором, влияющим на бизнес-процесс, необходимо выделять и оценивать эффект каждого изменения отдельно. Если же эффект невозможно разделить между инициативами, они оцениваются совместно как единый кейс (каскад моделей).

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

рост конверсии в кредиты одновременно сопровождался снижением ставки по рынку — эффект от ИИ-модели отклика необходимо изолировать от влияния ценового изменения

Ритейл и e-com

параллельно с внедрением рекомендательной системы был обновлён интерфейс каталога — эффект от ИИ оценивается отдельно через изолированный А/Б тест

Транспорт и логистика

снижение времени доставки достигалось одновременно за счёт ИИ-маршрутизации и расширения автопарка — вклад каждого изменения рассчитывается отдельно

АПК, энергетика и промышленность

сокращение простоев оборудования достигнуто и заменой части узлов, и внедрением предиктивной модели — необходимо разделить финансовые эффекты каждого фактора

ИКТ

отток снизился после одновременного запуска модели удержания и акционного предложения — эффект ИИ выделяется через контрольную группу без акции

05 Принцип полноты

Необходимо учитывать **все** инкрементальные расходы и доходы, связанные с достижением бизнес-цели: по целевому продукту и по смежным (каннибализация и кросс-продажи); по целевому каналу и по всем доступным каналам; прямые и косвенные расходы всех связанных подразделений; текущие расходы и расходы на поддержку модели в будущем.

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

ИИ-модель повышает конверсию в потребительские кредиты, но одновременно снижает продажи кредитных карт и страховых продуктов — каннибализация по всему продуктовому сплиту учитывается в итоговом эффекте

Ритейл и e-com

рост онлайн-продаж через рекомендательную систему снижает продажи в офлайн-магазинах сети — необходимо учитывать межканальный переток

Транспорт и логистика

оптимизация маршрутов снижает затраты на топливо, но увеличивает нагрузку на диспетчеров — оба эффекта включаются в оценку

АПК, энергетика и промышленность

снижение брака за счёт ИИ-контроля качества уменьшает затраты на утилизацию и переработку, но требует дополнительного обслуживания видеосенсоров, ИТ-инфраструктуры и затрат рабочего времени на обработку ложноположительных срабатываний — все эти инкрементальные расходы учитываются

ИКТ

модель оттока удерживает клиентов с помощью скидок — в оценку включается не только удержанный ARPU, но и стоимость предоставленных скидок

06 Принцип материальности

Нет необходимости стремиться рассчитать эффект с точностью до копейки. Необходимо ответить на вопрос: **насколько чувствительна прибыль к изменению тех драйверов, точность которых мы хотим повысить?** Если уточнение оценки не влияет на итоговое решение об экономической целесообразности инициативы — дополнительные ресурсы на повышение точности расчёта не нужны.

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

уточнение ставки дисконтирования с 15,0% до 15,1% не меняет вывод о целесообразности ИИ-инициативы — такое уточнение нецелесообразно

Ритейл и e-com

детализация эффекта кросс-продаж по каждому из 300 SKU избыточна, если суммарный кросс-продажный эффект составляет менее 1% от общего эффекта инициативы

Транспорт и логистика

точная аллокация расходов на сопровождение модели между 12 региональными подразделениями не влияет на общий ROI — достаточно расчёта на уровне дивизиона

АПК, энергетика и промышленность

стоимость одного сбоя оборудования оценивается укрупнённо — детальный расчёт по каждому типу узла нецелесообразен, если итоговая оценка не меняет решения о внедрении

ИКТ

при оценке оттока учитываются клиенты с ARPU выше порогового значения — сегменты с минимальной выручкой исключаются как несущественные

07 Принцип учета консолидированного эффекта

Оценку необходимо проводить **на уровне всей группы компаний / холдинга** (если применимо), чтобы учесть все связанные финансовые и налоговые последствия. Эффект, зафиксированный на уровне одного юридического лица, может быть нейтрализован или усилен на уровне группы.

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

ИИ-модель повышает маржинальность по одному продукту внутри банка, но снижает кросс-продажи смежных продуктов в другом юридическом лице группы — итоговый эффект оценивается на уровне консолидированного P&L группы

Ритейл и e-com

рекомендательная система интернет-магазина перераспределяет продажи между юридическими лицами группы с разными налоговыми режимами — требуется консолидированный расчёт

Транспорт и логистика

экономия топлива в одном дочернем предприятии транспортного холдинга компенсируется ростом затрат на обслуживание парка в другом — оценивается суммарный эффект по группе

АПК, энергетика и промышленность

ИИ-оптимизация производства на одном заводе холдинга перераспределяет загрузку на другие предприятия — учитывается совокупное влияние на операционные расходы группы

ИКТ

снижение оттока в одном сегменте за счёт агрессивных скидок может каннибализировать доходы другого сегмента — важно смотреть на итог в целом

Объект оценки: AS IS/TO BE

Одно из основных заблуждений, которое может быть: если модель валидирована, имеет отличные метрики качества и работает, то значит, что все отлично.

Цель внедрения моделей ИИ различается в зависимости от организации, однако с точки зрения финансовой оценки важно понимать, что успешное влияние внедренной ИИ модели на PnL или любую другую денежную

метрику года не всегда связано с метрикой качества этой модели. Одно из основных заблуждений, которое может быть, это ситуация, при которой единственными критериями успешности модели являются отличные метрики качества. Важно смотреть на процесс внедрения моделей ИИ широко и учитывать все потенциальные эффекты в соответствии с принципом полноты оценки.

— ПРИМЕРЫ

Финансовая индустрия

Модель максимизирует продажи потребительских кредитов из всей продуктовой полки. Эффект — это только прирост прибыли по кредитам? Нет: если на той же полке есть кредитные карты, страховки и другие продукты, возникает каннибализация с учётом ограниченного времени продающих сотрудников и замещения продуктов-аналогов. Итоговый эффект нужно считать по всему сплиту продаж.

Ритейл и e-com

Модель персонализации увеличивает конверсию в покупку онлайн. Означает ли это рост выручки компании в целом? Не обязательно: если часть покупок мигрирует из офлайн-канала в онлайн без роста общего числа транзакций, реальный инкрементальный эффект окажется значительно ниже, чем показывает конверсия онлайн-канала в изоляции.

Транспорт и логистика

Модель оптимизирует маршруты и снижает пробег. Финансовый эффект — только экономия топлива? Нет: необходимо также учитывать изменение нагрузки на водителей, время доставки (влияет на штрафы и лояльность клиентов) и возможный рост количества рейсов за счёт высвободившегося времени.

АПК, энергетика и промышленность

Модель предиктивного обслуживания снижает количество аварийных остановок оборудования. Это снижение простоев — и есть весь эффект? Нет: аварийный ремонт также влечёт за собой брак продукции, задержки поставок и штрафные санкции со стороны покупателей — все эти компоненты должны быть учтены в оценке.

ИКТ

Голосовой/чат-бот в контактном центре сокращает время обработки обращений. Эффект — только экономия на операторах? Нет: нужно учитывать влияние на удовлетворённость и отток (если бот ухудшает клиентский опыт, рост оттока съедает экономию), долю недозвонов и повторных обращений, а также нагрузку на оставшихся операторов по сложным запросам.

Реальность очень сложна, поэтому важно заглянуть глубже - внутрь процесса - и учесть взаимосвязи между ролями, продуктами и компонентами прибыли. Таким образом, мы приходим к пониманию, что:

Объект финансовой оценки - изменения процессов с применением ИИ модели.

Задача финансовой оценки - оценить прямые и косвенные последствия изменения или создания процессов.

Тогда нам нужно ответить на следующие вопросы:

- 1 Как процесс устроен AS IS (бизнес-процесс без ИИ воздействия)?
- 2 Какой драйвер в результате будет изменяться?
- 3 Как мы будем собирать статистику по процессу до и после внедрения модели, как будет выглядеть процесс TO BE (бизнес-процесс с ИИ воздействием), какие могут быть риски?
- 4 Какие будут эффекты, финансовые и нефинансовые, если будет внедрена модель?

- 5 Влияет ли изменение только на 1 объект управления (продукт, услугу, процесс) или же и на иные объекты управления, на которые не направлено воздействие напрямую?

Инкремент

Инкремент — изменение драйвера, метрики, финансового эффекта, возникшее в результате конкретного воздействия на процесс по итогу внедрения конкретной модели ИИ / инициативы с ИИ.

Понятие инкремента неразрывно связано с понятием **бейзлайна** — процесса или состояния процесса, относительно которого измеряется инкремент, — а также с **Принципом 4 «Связь с инициативой»**.

Важно понимать, что бейзлайн для финансовой оценки / для оценки экономического эффекта от ИИ инициативы не всегда равен процессу AS IS. Бейзлайн для оценки представляет собой лучшую ближайшую альтернативу процессу TO BE.

- Если процесс AS IS представляет собой лучшую ближайшую альтернативу, то для получения оценки достаточно сравнить процессы AS IS и TO BE между собой
- Если создается новый продукт / процесс – анализируются альтернативные издержки ближайшей альтернативы (например, Business Intelligence (BI) решение) для достижения аналогичной цели, и оценивается ИИ инкремент относительно выработанных BI инструментов. Тут необходимо также руководствоваться принципом материальности и не создавать слишком дорогую, но точную BI альтернативу.

Крайне важно понимать: эффект от внедрения ИИ-модели измеряется именно против лучшей альтернативы. Иначе в оценку дохода от ИИ неизбежно попадёт доход от самого факта автоматизации — без вклада ИИ.

Выбор бейзлайна: примеры по отраслям

— ПРИМЕРЫ ПО ОТРАСЛЯМ

Финансовая индустрия

Исходящий обзвон клиентов ИИ-ботом для напоминания о брошенной заявке. Бейзлайн — не «без звонка», а лучший из доступных вариантов: звонок оператором или звонок с предзаписанной аудиодорожкой.

Ритейл / e-com

Система динамического ценообразования. Бейзлайн — не «без ценообразования», а лучшая действующая практика: таблица сезонных скидок, управляемая аналитиком, или простые правила наценки.

ИКТ

Модель предсказания оттока. Бейзлайн — существующий набор BI-правил сегментации «тихих» абонентов, который использовался для проактивных предложений до внедрения ИИ.

Производство

ИИ-система предиктивного обслуживания оборудования. Бейзлайн — регламентное обслуживание по фиксированному расписанию (текущий AS IS процесс), так как более умной альтернативы до внедрения модели не существовало.

Транспорт и логистика

ИИ-система управления доходностью рейсов. Бейзлайн — не «продажа билетов по фиксированной цене», а лучший из доступных вариантов: классическая тарифная сетка, которой специалист управляет вручную по простым правилам.

Проблема бейзлайна в сложносоставных и долгосрочных проектах

Трудности с бейзлайном возникают в двух типичных ситуациях.

1. Составные инициативы. Когда одновременно создаётся новый процесс и внедряется ИИ, оценка инкремента только от ИИ требует явного разделения эффектов (см. Принцип 4).

2. Долгосрочные проекты с итеративным обновлением модели. Базовой рекомендацией методологии является расчёт полного экономического эффекта относительно лучшей альтернативы без ML. Однако на горизонте нескольких лет поддерживать изначальный бейзлайн становится сложно. В таких случаях часто прибегают к оценке относительно предыдущей версии модели, а итоговый экономический эффект считают как сумму инкрементальных изменений.

Важно: сумма инкрементальных изменений не всегда равна совокупному эффекту относительно бейзлайна, так как может присутствовать каннибализация (или, реже, синергия) между несколькими версиями обновлений:

Инкремент (V1 vs V3) ≠ Инкремент (V1 vs V2) + Инкремент (V2 vs V3)

При подходе с суммированием инкрементов рекомендуется **раз в год** проводить «обратный» эксперимент — замер эффекта последней версии модели относительно версии, существовавшей на начало года (или на k поколений назад) — и использовать эту калибровку для корректировки суммы инкрементов.

Мониторинг эффектов

Чтобы следовать принципу измеримости, а также иметь возможность в любой момент жизни ИИ инициативы ответить на вопрос: какую экономическую и финансовую выгоду она приносит, – в рамках процесса оценки необходимо всегда выставлять **регулярный мониторинг** оцененного эффекта.

Регулярный мониторинг выставляется как на метрики качества модели (Precision, Recall, GINI, MAPE, MAE, т.д.), так и на финансовые и операционные драйверы оценки.

Метрики качества

Все модели деградируют со временем, если их не дообучать, из-за изменения входного потока данных и изменения закономерностей.

Для разного типа моделей мониторятся разные метрики точности. Какие именно - зависит от типа ML задачи и критичности бизнес-процесса.

Как правило, чем выше цена ошибки при работе модели (например, выдача кредита, блокировка транзакции) тем ниже толерантность к допустимому снижению метрик модели.

Модель обновляется, когда мы видим снижение метрик ниже допустимого порога.

Перспективным выглядит автомониторинг моделей; в отсутствие автомониторинга – мониторинг качества производится с определенной регулярностью в зависимости от критичности процесса, например модели высокой степени значимости мониторятся чаще, чем модели низкой степени значимости.

Рекомендуемая частота замеров качества модели – раз в месяц.

Минимально допустимая в зависимости от степени значимости моделей:

КРИТИЧЕСКИ ЗНАЧИМЫЕ
(НАПР., ЗНАЧИТЕЛЬНЫЙ
ФИН. ЭФФЕКТ ИЛИ
ЗАТРАГИВАЕТ
КРИТИЧЕСКИ ЗНАЧИМЫЕ
ПРОЦЕССЫ)

ВЫСОКОЙ СТЕПЕНИ
ЗНАЧИМОСТИ
(НАПР., СРЕДНИЙ
ФИН. ЭФФЕКТ ИЛИ
ЗАТРАГИВАЕТ
ЗНАЧИМЫЕ
ПРОЦЕССЫ)

СРЕДНЕЙ СТЕПЕНИ
ЗНАЧИМОСТИ
(НАПР., НИЗКИЙ ФИН.
ЭФФЕКТ ИЛИ
ЗАТРАГИВАЕТ
ПРОЦЕССЫ СРЕДНЕЙ
ЗНАЧИМОСТИ)

НИЗКОЙ СТЕПЕНИ
ЗНАЧИМОСТИ (НАПР., БЕЗ ФИН.
ЭФФЕКТА И НЕ ИМЕЮЩИЕ
СУЩЕСТВЕННОГО ВЛИЯНИЯ НА
ВАЖНЫЕ ПРОЦЕССЫ)

Раз в 3 месяца

Раз в полгода

Раз в год

Не проводится

Финансовые и операционные метрики

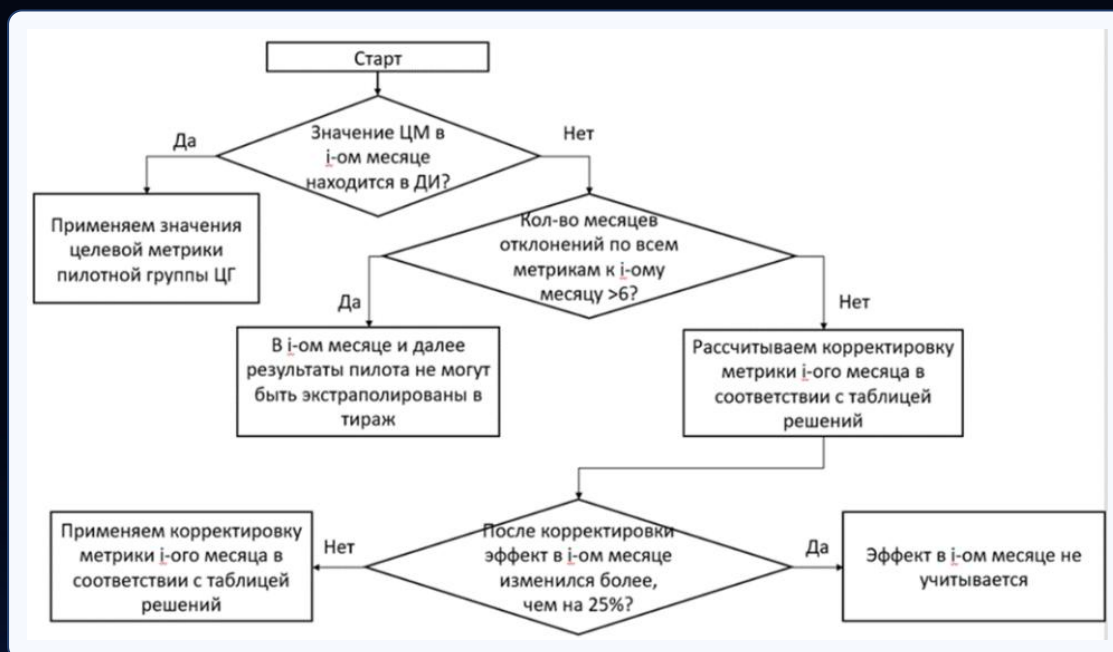
Что касается финансовых и операционных метрик, частота и метод мониторинга сильно зависят от типа эффекта рассматриваемой ИИ инициативы, а также наличия или отсутствия постоянной контрольной группы.

Для случая, когда постоянная КГ (СКГ) присутствует, необходимо проводить замер инкремента драйверов относительно этой контрольной групп; при этом необходимо проверить СКГ на предмет вымывания/эрозии.

- Не реже раз в 6 месяцев с момента первой оценки для эффектов на рост дохода, снижение операционных расходов, снижение операционных рисков и высвобождение ликвидности рекомендуется проводить мониторинг.
- Не реже раза в 12 месяцев для эффектов на высвобождение штатной численности, восстановление/недосоздание резервов, а также снижения расходов на капитал рекомендуется проводить мониторинг в конце каждого календарного/финансового года.

Для случая, когда постоянная КГ (СКГ) отсутствует, рекомендуется проводить ежемесячную проверку эффекта и сравнивать его с доверительным интервалом, полученным на этапе проведения А/Б теста.

На примере:



Чистый vs Валовый эффект

В оценку ИИ-проектов необходимо включать все затраты на создание и поддержку. Считаем, что чистый (net) подход самый правильный для расчета фин. эффекта от ИИ.

Преимущества чистого подхода:

- Полный учет инвестиций и реалистичная оценка прибыльности (ROI)
- Объективное сравнение проектов
- Эффективное управление рисками

Учитываемые затраты.

При расчете чистого эффекта учитываются все расходы, относимые к созданию и поддержке ИИ:

1 Фонд оплаты труда (ФОТ):

1. Все расходы связанные с человеческим капиталом:
 - валовый оклад, налоги (помимо тех, что учтены в валовом окладе), соц взносы, все компенсации, районные надбавки, отпускные, бонусы, премии
 - вмененные расходы (аллокация стоимости офисного места, техники используемой сотрудником); необходимость учета вмененных расходов определяется финансовой службой на основании предоставленных обоснований.
2. Оценивается по PnL (амортизация нематериального актива), не с точки зрения движения денежных средств;
3. При оценке важно уточнить учётную политику конкретной компании: в ряде случаев затраты на разработку ИИ капитализируются как НМА, однако чаще — особенно в быстро меняющихся отраслях — они относятся на расходы периода из-за сложности определения СПИ;
4. Аллокация сотрудников: если сотрудник был привлечен в разработку ИИ-продукта или процесса, то затраты на него учитываются в затратах на ИИ.

2 Серверная инфраструктура:

1. Расходы учитываются с позиции PnL отчета, а не с точки зрения движения денежных средств;
2. Делятся на 2 категории - инфраструктура, закупленная только под инициативы ИИ (например GPU), и общебанковская инфраструктура, которой пользуются команды для создания ИИ продуктов на основе внутренних тарифных моделей;
3. Первая категория полностью относится к "инвестиции в ИИ". В расход берутся амортизация серверов, часть расходов датацентров, аллоцированных на данные сервера, команду (тех поддержка оборудования / сети), необходимые лицензии для поддержания данных серверов, если есть;
4. Вторая категория аллоцируется на сущность "инвестиции в ИИ" пропорцией из метрик потребления общебанковских платформ (платформа виртуализации, базы данных, логирование, аналитические платформы, инференс и тд). Важно отметить, что это аллокация фактических затрат, а не биллинг. Из-за этого стоимость единицы ресурса не фиксированная, а зависит не только от общей стоимости платформы, но и от общего потребления / утилизации по компании инфраструктурных платформ.

3 Внешние расходы:

1. Все внешние закупки - лицензии / консалтинг / услуги / поддержки / разработка с использованием персонала вне штата/ внешняя (аутсорс) разработка;
2. Важно определить драйвер аллокации внешних расходов, например:
 - Приписывание на команду, запросившую расход
 - На основе использованных токенов/карт/иных счетных и отделимых параметров
 - На основе внутренних тарифных моделей

Гранулярность оценки каждого из пункта выше - 1 раз в месяц.

Метрики

Какими метриками можно оценить финансовый эффект от ИИ?

Типы финансовых эффектов

Можно выделить 4 универсальных типа финансовых эффектов от ИИ-инициатив, применимых к компаниям любой отрасли, а также 4 специфических типа, характерных для финансовой индустрии.

Универсальные типы финансовых эффектов

ТИП ЭФФЕКТА ОТ ИИ-ИНИЦИАТИВЫ	НА КАКУЮ ФИНАНСОВУЮ МЕТРИКУ ОКАЗЫВАЕТСЯ ВЛИЯНИЕ	РЕКОМЕНДУЕМЫЙ УРОВЕНЬ ОЦЕНКИ
Реальное влияние на P&L		
1) Рост доходов	Операционный доход	I, II, III
2) Снижение операционных расходов	Операционные расходы	I, II, III
3) Высвобождение штатной численности (ПШЕ)	Операционные расходы	I, II, III
4) Снижение операционных рисков	Операционный доход / Операционные расходы	I, II, III

Что такое эффект роста доходов?

Наиболее простой для понимания тип эффекта — рост доходов. Простота данного эффекта, тем не менее, требует внимательности при оценке прямых и косвенных последствий реализуемых изменений. Влияет на операционный доход компании.

Бизнес-драйверы роста дохода делятся на две группы:

- Изменение в **количестве продаж**
- Изменение в **маржинальности или прибыльности продаж**

Причём продажи и маржинальность могут рассматриваться как в разрезе отдельного продукта, так и в совокупности по нескольким продуктам в рамках продуктового сплита.

За счёт чего может изменяться цена и количество продаж:

- Повышение производительности / автоматизация процесса
- Запуск новых продуктов
- Изменение модели отклика в кампаниях продаж / рост эффективности коммуникаций с клиентами
- Изменение среднего чека / средней ставки / средней маржи от одной новой продажи

- Снижение текучести сотрудников продаж и повышение эффективности подбора персонала
- Рост транзакционной активности текущих клиентов
- Удержание клиентов
- Расширение клиентской базы

По данным [McKinsey \(2025\)](#), рост доходов от ИИ наиболее часто фиксируется в маркетинге и продажах, разработке продуктов и стратегическом планировании.

Что важно анализировать при эффекте роста дохода?

Необходимо анализировать побочные эффекты помимо прямого воздействия на целевой драйвер:

- **Эффект каннибализации** — снижение продаж конкурирующих продуктов и услуг при росте продаж целевого продукта, или снижение продаж целевого продукта в конкурирующем канале сбыта
- **Эффект кросс-продаж** — рост продаж нецелевых продуктов и услуг как сопутствующий эффект

Примеры по отраслям:

Финансовая индустрия. Модель отклика повышает конверсию в потребительские кредиты. При этом необходимо проверить весь продуктовый сплит: кредитные карты (возможная каннибализация), страховки (возможный кросс-продажный эффект), а также кросс-канальную конкуренцию между онлайн- и офлайн-продажами.

Ритейл и e-com. Рекомендательная система увеличивает средний чек. Одновременно нужно оценить, не происходит ли снижение частоты повторных визитов (клиент купил «всё сразу») и не каннибализируются ли продажи соседних категорий с более высокой маржой.

Транспорт и логистика. ИИ-система динамического ценообразования увеличивает выручку на маршруте. Важно проверить, не происходит ли отток клиентов в пользу конкурентов с фиксированными тарифами и не снижается ли загрузка в непиковые часы.

АПК, энергетика и промышленность. ИИ-система прогнозирования спроса снижает недопоставки и увеличивает выручку. Необходимо оценить, не возникает ли при этом избыток запасов по смежным позициям, который увеличивает затраты на хранение.

ИКТ. Модель персонализированных предложений увеличивает ARPU целевого сегмента. Нужно проверить, не снижает ли это конверсию в более дорогие тарифы у клиентов, которые без модели перешли бы на более высокий план.

Что такое срок жизни продукта? Зачем его учитывать в эффекте роста доходов?

В расчёте эффекта роста дохода необходимо всегда учитывать срок жизни продаваемых продуктов и среднюю прибыль одного проданного продукта. Если по одному продукту срок жизни 1 месяц, а по другому — 10 лет, то при одновременном росте конверсии по первому и падении по второму общий эффект может оказаться отрицательным — даже если прирост конверсии по первому продукту выглядит более значимым. Расчёт в таком случае строится **повинтажно** — с учётом момента возникновения продаж и срока жизни каждого продукта.

В качестве альтернативы продуктовому подходу можно использовать метрику CLTV (Customer Lifetime Value), если она доступна и корректно рассчитывает ценность клиента. CLTV применима для внутренней оценки; для внешней отчётности рекомендуется использовать метрики в терминах P&L.

Что такое эффект снижения операционных расходов?

Эффект от снижения операционных расходов аналогичен росту доходов, только с обратным знаком: чем больше сокращаются расходы, тем больший эффект наблюдается. Драйверами эффекта являются количество совершаемых расходных операций и их стоимость.

По данным [McKinsey \(2024\)](#), наибольшую долю компаний, фиксирующих снижение затрат от ИИ, показывают операционное обслуживание и HR-функции. BCG подтверждает: ряд крупных внедрений обеспечивает 8–12% снижение затрат по всей цепочке создания стоимости.

Примеры по отраслям:

Финансовая индустрия. Внедрение ИИ-маршрутизации в колл-центре снижает среднее время обработки обращения (АНТ) и объём входящих звонков. BCG фиксирует снижение объёма звонков на ~35% и сокращение АНТ на ~20% при внедрении ИИ в поддержку клиентов.

Ритейл и e-com. ИИ-система управления запасами сокращает долю неликвидных остатков и расходы на хранение. BCG фиксирует снижение запасов на \$100 млн и улучшение точности прогнозирования на 15 п.п. при внедрении ИИ в supply chain.

Транспорт и логистика. ИИ-маршрутизация курьеров оптимизирует затраты на топливо за счёт сокращения пробега вхолостую. Драйвер — километраж на единицу доставки; мониторинг ведётся по телематическим данным.

АПК, энергетика и промышленность. ИИ-контроль качества продукции на производственной линии снижает долю брака и затраты на утилизацию и переработку. Драйвер — стоимость одного дефектного изделия и доля брака в общем объёме выпуска.

ИКТ. ИИ-автоматизация заявок обрабатываемых операторами второй линии технической поддержки с ИИ-ассистентами. Драйвер — доля заявок, закрытых удалённо, и стоимость одного выезда.

Что такое эффект высвобождения штатной численности (эффект ПШЕ)? Какие примеры эффектов ПШЕ есть?

Эффект от высвобождения штатной численности измеряется в терминах количества сотрудников / ПШЕ (полных штатных единиц). Ключевые драйверы — количество совершаемых операций и их нормативная длительность.

Высвобождение численности подразделяется на два типа:

- **Фактические (реальные) сокращения сотрудников**
- **Исключение потребности в дополнительных сотрудниках (нenaйм)**

***Важно:** эффективность ИИ в сравнении с «ручным» трудом должна быть доказана через пилотирование или хронометраж. Расчёт должен быть подтверждён владельцем моделей численности и найти отражение в бизнес-плане компании — иначе эффект останется виртуальным.*

Примеры по отраслям:

Финансовая индустрия. Внедрение ИИ-модели в процесс принятия решений по кредитным заявкам повышает коэффициент автоматизации с 60% до 90%. При росте входящего потока заявок вдвое фактическая численность операторов сокращается с 4 до 2 ПШЕ (реальное сокращение), и одновременно исключается потребность в найме 4 дополнительных сотрудников (эффект ненайма).

Ритейл и e-commerce. ИИ-система автоматической разметки и категоризации товаров исключает необходимость расширения команды контент-менеджеров при росте ассортимента. Эффект ненайма рассчитывается через норматив: количество новых SKU в год / производительность одного контент-менеджера.

Транспорт и логистика. ИИ-планировщик рейсов исключает потребность в расширении диспетчерского штата при росте объёма перевозок. Эффект подтверждается сравнением нормативной численности диспетчеров при ручном и автоматизированном планировании.

АПК, энергетика и промышленность. ИИ-визуальный контроль качества на конвейере замещает часть ручных инспекторов. Реальное сокращение ПШЕ подтверждается хронометражем операций до и после внедрения, при этом обязательно учитываются нормативные ограничения по минимально допустимому составу дежурной смены согласно правилам промышленной безопасности

ИКТ. ИИ-чат-бот первой линии поддержки снижает нагрузку на операторов контакт-центра. При росте клиентской базы компания исключает необходимость найма новых операторов — эффект рассчитывается через среднее количество обращений на оператора в месяц.

Эффект реального сокращения численности

Данный тип эффекта подразумевает воспроизведение части функционала сотрудников за счет ИИ алгоритмов, в результате чего происходят автоматизация функционала и оптимизация нормативной численности, т.е. сокращаем реальное количество сотрудников.

При расчете финансового эффекта от фактического сокращения, помимо очевидного прямого эффекта от сокращения расходов на заработную плату, нужно учесть все сопутствующие расходы и доходы, которые свойственны рассматриваемому процессу. Например, необходимо учесть выплаты сотруднику дополнительных денежных средств в случае досрочного прекращения трудового договора со стороны работодателя.

При переходе от ручного труда к ИИ решениям могут возникать эффекты на иные показатели эффективности оптимизируемого процесса, и эта связь с ключевым драйвером, улучшаемой метрикой, не всегда очевидна, но требует внимания, следуя принципу полноты

Например, инициатива автоматизации принятия решений может уменьшить количество операционных дефектов, стоимость которых отразится на PnL. Возможен сценарий, когда автоматизация может привести и к увеличению объемов продаж или уменьшению операционных издержек, или же ухудшить эффективность продаж при полной автоматизации функционала. Многие эффекты, связанные с замещением ручного труда, являются «многослойным», с большим количеством драйверов внутри бизнес процесса, которые необходимо измерять до и после изменений и мониторить в динамике.

Эффект исключения потребности в дополнительных сотрудниках (ненайм)

Эффект возникает, когда при росте бизнес-драйвера численности (росте количества клиентов, которых нужно обслуживать, росте заявок, которые нужно обрабатывать, и т.д.) увеличения численности не происходит за счет внедрения ИИ модели. Такого рода модель может, например, увеличивать эффективность работы текущего штата сотрудников, что позволяет им обрабатывать и принимать решения на большем объеме входящего потока задач с меньшими издержками по времени.

Важным моментом в данном случае является фокус на том, как считается эффект ненайма – это должно быть объективной оценкой на основании драйверов процесса, количества выполняемых операций и их длительности. При этом функционал должен быть нормативно обязательным для подразделения.

Рассмотрим конкретный пример, покрывающий одновременно эффект реального сокращения ПШЕ и эффект ненайма: в качестве бейзлайна инициативы используются ИИ модели, внедренные в процесс принятия решения по кредитным заявкам. Для проверки того, что автоматизация проводилась именно за счет ИИ, была сделана кросс-проверка функций, выполняемых штатной численностью, и драйверов процесса, на которые влияли указанные ИИ модели.

За счет внедрения ИИ, коэффициент автоматизации обработки заявок был повышен с 60% до 90%. В начале входящий поток заявок был 100 шт. 40% заявок обрабатывалось в ручном режиме, на что требовалось 4 сотрудника (производительность – 10 заявок на человека).

год	0	1
Коэффициент автоматизации	60%	90%
Количество входящих заявок для обработки	100	200
Численность ПШЕ	4	2
Эффекты:		
Реальные сокращения		2
Ненайм		4

При одновременно повышении входящего потока заявок со 100 штук до 200 и коэффициента автоматизации с 60% до 90% в следующем году потребовалось 2 сотрудника на ручную обработку. Размер эффекта реального сокращения и ненайма в терминах ПШЕ:

- Реальное сокращение посчитать довольно просто: поскольку мы знаем (проверили это заранее), что на процесс не влияло ничто, кроме рассматриваемой ИИ инициативы, фактические сокращения равны численность на начало периода минус численность на конец периода. Наблюдается эффект в 2 ПШЕ.
- Ненайм в такой конфигурации рассчитывается несколько сложнее. Представим, что ИИ модель не внедряли бы, при этом на процесс не оказывалось никакого стороннего влияния (единственное, что могло меняться от года к году – это улучшение ИИ модели, дообучение/включение новых параметров). В таком случае, коэффициент автоматизации в последующих годах остался бы на уровне 60%, или потребовалось бы 8 человек для ручных проверок, 4 из которых уже работали изначально. Таким образом, пришлось бы нанять дополнительно 4 сотрудника в штат. Эффект ненайма получается 4 ПШЕ.

Что такое эффект на снижение операционных рисков, приводящих к недополучению прибыли / росту операционных расходов?

Эффект от снижения операционных рисков влияет на P&L через два основных драйвера:

- **Операционные дефекты** — ошибки в бизнес-процессах, стоимость которых отражается в расходах или недополученных доходах
- **Фрод / мошенничество** — потери от недобросовестных действий внешних или внутренних участников

По обоим драйверам необходимо либо вести статистику для оценки среднего «вклада» в P&L, либо проводить А/Б тест с контролем всех изменяемых переменных. По выбранным драйверам обязателен постоянный мониторинг: если модель год к году выявляет всё больше фрода/дефектов при стабильном или снижающемся их общем количестве — это сигнал для проверки деградации алгоритма.

Примеры по отраслям:

Финансовая индустрия. ИИ-модель прогнозирует вероятность операционного дефекта сотрудника банковского офиса. Эффект двойной: (1) высвобождение аудиторов за счёт более точечных проверок, (2) сокращение издержек от дефектов, статистика стоимости которых ведётся системно.

Ритейл и e-commerce. ИИ-модель выявляет мошеннические возвраты товаров (фиктивный возврат, подмена товара). Драйвер эффекта — доля выявленного фрода в общем объёме возвратов и средняя стоимость одного случая.

Транспорт и логистика. ИИ-мониторинг поведения водителей снижает количество ДТП и штрафных ситуаций. Драйвер — стоимость одного инцидента (ремонт, штрафы, простой) и частота инцидентов до и после внедрения.

АПК, энергетика и промышленность. ИИ-предиктивное обслуживание снижает вероятность аварийной остановки оборудования. Драйвер — стоимость одного часа простоя производственной линии и частота аварийных остановок, а также сопутствующие прямые потери (например, порча сырья в процессе обработки или повреждение смежных узлов при экстренной остановке)

ИКТ. ИИ-система выявления аномального трафика снижает потери от фрода при международных звонках. Драйвер — объём фродового трафика (минуты/руб.) и скорость его выявления и блокировки.

Финансовая индустрия

Следующие типы эффектов применимы исключительно к организациям финансового сектора (банки, страховые компании, управляющие компании) ввиду специфики регуляторного учёта, управления капиталом и ликвидностью. Для компаний других отраслей данный раздел носит ознакомительный характер.

ТИП ЭФФЕКТА	НА КАКУЮ ФИНАНСОВУЮ МЕТРИКУ ОКАЗЫВАЕТСЯ ВЛИЯНИЕ	РЕКОМЕНДУЕМЫЙ УРОВЕНЬ ОЦЕНКИ
Реальное влияние на P&L		
5) Высвобождение ликвидности	Операционный доход	II, III
6) Восстановление / недосоздание резервов (фактическое влияние на кредитный риск)	Расходы на резервы	II, III
Косвенное влияние на P&L		
7) Высвобождение резервов (без влияния на фактический кредитный риск) — учётный эффект	Расходы на резервы в отчётный период с последующим восстановлением на сроке жизни продукта	III
8) Снижение расходов на капитал (за счёт снижения чистых затрат на капитал, в т.ч. от экономии на RWA)	Стоимость капитала	III

Что такое эффект на высвобождение ликвидности?

Данный тип эффекта возникает, когда ИИ модель высвобождает объем денежных средств, ликвидности, которую нужно держать, например, в рамках норматива ликвидности, или СДО денежных средств в кассовых центрах или банкоматах. Высвобождаемую ликвидность можно размещать по ставке Overnight, что обеспечит дополнительную прибыль компании

Драйвером эффекта в данном случае будет выступать объем ликвидности, на который влияет ИИ. Ставка, под которую размещаются денежные средства, является также изменяемым фактором, но она экзогенна по отношению к ИИ инициативе и нужна для мониторинга эффекта, поскольку от нее зависит доходность. Важна она еще и потому что, если единственной альтернативой размещения высвобожденной ликвидности будет размещение ее под отрицательную ставку, не важно, какой объем нам позволяет сэкономить ИИ, финансовый эффект такого действия будет отрицательным.

Также важно анализировать ошибки первого и второго рода в решениях, связанных с управлением ликвидностью: если мы ошибемся с оценкой свободной ликвидности для размещения ее на рынке, может реализоваться риск неспособности выполнить договорные отношения с клиентами из-за дефицита ликвидности, в результате чего компании придется привлекать средства по более высокой ставке на рынке.

Что такое эффекты на резервы? Что важно знать при оценке этих эффектов?

При расчете эффекта резервов необходимо корректно определить тип резервов, драйверы изменения резервов, понять, постоянный или временный эффект мы имеем, а также учесть принцип неттирования.

Важно помнить, что существует два типа резервов:

- 1) Первый тип – это регуляторные резервы. Это пруденциальные резервы, устанавливаемые Банком России, которые коммерческий Банк обязан перечислять на корреспондентский счет в Банке России.

Согласно российскому законодательству, Банки должны формировать Обязательные Резервы, Резервы под возможные потери по ссудам, Резервы под возможные потери и Резервы под обесценение ценных бумаг. Первые три типа наиболее часто встречаются при оценке ИИ эффектов. Кратко про каждый из них.

- Обязательные резервы - это часть суммы депозитов Банка, которую необходимо хранить на счете в Банке России.
- Резервы на возможные потери по ссудам (их сокращенно называют РВПС) формируются при потере ссудной стоимости, из-за ненадлежащего исполнения обязательств заемщиком. Например, клиент, которому был выдан кредит, задерживает очередной платеж на 40 дней. Банку необходимо сформировать дополнительный резерв по такой ссуде.
- Резервы по возможным потерям (сокращенно – РВП) формируются по балансовым активам с риском потерь (за исключением тех активов, по которым сформированы РВПС), по условным обязательствам и прочим потерям

2) Второй тип резервов – резервы по стандартам отчетности, простыми словами, это фонд, создаваемый организацией для покрытия своих убытков и отражаемый в отчетности. Два основных стандарта отчетности: международные стандарты финансовой отчетности (МСФО) и российские стандарты (РСБУ). Резервы по МСФО и по РСБУ не равны, так как в методологиях содержатся различные правила учета отдельных статей баланса.

Можно выделить три типа эффектов на резервы:

- временное сокращение/увеличение резервов;
- увеличение/уменьшение срока сбора просроченной задолженности;
- изменение объема сбора просроченной задолженности.

Давайте разберем каждый из типов подробнее на примерах.

1) Временное сокращение/увеличение резервов

Пусть разработанный ИИ-алгоритм проверки базы данных просроченных кредитов показал, что начисление резерва в размере 25% по ряду займов на сумму 400 млн руб. не требуется, поскольку они обеспечены залогом. По таким займам норма резервирования, например, должна быть 0%. То есть мы начисляем резерв 25%, а можно начислять 0%. И мы исправляем эту ошибку.

Исправление ошибок позволило сократить резервы на 100 млн руб. ($400 \times 25\%$). Пусть ошибочно созданные резервы были бы восстановлены без модели через 180 дней. 180 дней - это длительность эффекта, мы имеем временный эффект. Эффект, учитывая стоимость денег во времени, составит 2,5 млн руб.

При подтверждении таких эффектов надо помнить, что эффект не признается по кейсам, когда изменяется политика Банка по управлению резервами или нормы регулятора по изменению учета РВПС. Так как изменение резервов в таких случаях не является последствием внедрения ИИ-моделей. Модель в данном случае нужна, чтобы выполнить изменившиеся обязательные требования.

2) Увеличение/уменьшение срока сбора просроченной задолженности

Ускорение процесса сбора проблемной задолженности приводит к более раннему поступлению денежных средств в Банк.

Пусть ИИ-алгоритм позволил сократить срок взыскания проблемной задолженности на 1 месяц. Объем взыскиваемой проблемной задолженности – 8 млн руб., покрытие резервами - 100%. Если средняя стоимость капитала составляет 20%, то эффект за год составит 133 тыс. руб. ($8 \times 0,2 = 1,6$ млн. $1,6$ млн./12 мес. = 133 тыс. за 1 месяц)

3) Изменение объема сбора просроченной задолженности

При увеличении сбора безнадежной задолженности эффект представляет собой всю сумму дополнительно собранной задолженности (не забываем, что за вычетом лучшего альтернативного варианта работы с безнадежной задолженностью).

Например, коллекторы в результате внедрения модели научились собирать проблемную задолженность, которая ранее признавалась безнадежной и продавалась по цессии за 10% номинала. Объем такой дополнительно собранной задолженности – 1 млрд руб., покрытие резервами – 100%. Поскольку взыскание такой безнадежной задолженности прямо попадает в капитал Банка, то эффект составит 900 млн руб.

Принцип неттирования в эффектах на резервы и почему его важно учитывать в эффектах на резервы.

Эффекты изменения резервов специфичны своим изменением во времени. В один период времени резерв формируется, что отражается в отчетности увеличением резерва, в другой период времени происходит высвобождение резервов, что также находит отражение в отчетности. Поэтому нужно смотреть life-time эффект, не ограничиваясь одним отчетным периодом. Для некоторых случаев влияния на резервы, без прямого влияния на кредитный риск, life-time эффект может вырождаться в 0. Такой случай был назван принципом неттирования.

Принцип неттирования для понимания лучше разобрать на примере.

Пусть есть ИИ-инициатива, внедрение которой приводит к тому, что меняется временная структура создания и высвобождения резервов. Например, модель рекомендует привлекать на продукт «Кредит» клиентов, отличных от тех, которые привлекались бы в процессе без ИИ. Для простоты будем рассматривать только клиентов, которые приобрели кредит в 2020 году.

Клиенты, привлекаемые с помощью ИИ (это процесс TO BE), оцениваются как более благонадежные, и по ним формируются резервы в размере 20 млн. руб., в то время как по клиентам, привлекаемым без ИИ (это процесс

AS IS), сформированы резервы размером в 32 млн. руб. Положительный эффект составил 12 млн. руб.

В 2021 году резервы снизились в AS IS процессе на 12,8 млн. руб. до 19,2 млн. руб., в TO BE – на 8 млн. руб. до 12 млн. руб. Возник отрицательный эффект в PnL 2021 года равный -4,8 млн. руб.

В 2022 году эффект в данном кейсе также отрицательный.

Таким образом, эффект резервов от года к году может меняться не только в объеме, но и по знаку (с + на – и наоборот), поэтому важно рассматривать такие эффекты life-time.

	ДАТА ОТЧЕТА		
	31.12.2020	31.12.2021	31.12.2022
Портфель 2020 на дату отчета	100	60	0
Резерв до	32	19,2	0
Резерв после	20	12,00	0
Дельта Резервы	12	7,2	0
Финрез Год-Год AS IS		12,8	19,2
Финрез Год-Год TO BE	12	8	12
Дельта финрез	12	-4,8	-7,2

Что такое эффект снижения расходов на капитал?

Говоря про эффект изменения расходов на капитал, стоит начать с того, что расходы на капитал состоят из трех элементов: чистые затраты на капитал, называемые ЧЗК, затраты капитала на резервы, и затраты на фондирование.

Поэтому эффект в этой части может возникать через изменение Чистых Затрат на Капитал (ЧЗК), через изменение затрат капитала на резервы и через изменение активов, взвешенных с учетом риска (RWA).

Рассмотрим пример:

Пусть есть ИИ-инициатива, внедрение которой приводит к тому, что меняется временная структура создания и высвобождения резервов. Например, модель рекомендует привлекать на продукт «Кредит» клиентов, отличных от тех, которые привлекались бы в процессе без ИИ.

Предположим, что в процессе TO BE клиенты имели значительные задержки в совершении платежей, и в Банке наблюдался высокий уровень просроченной задолженности.

Это значит, что в процессах AS IS и TO BE отличаются значения показателя RWA – активов, взвешенных с учетом риска.

Теперь разберемся с показателем ЧЗК. Чистые затраты на капитал - это объем задолженности с учетом процентов, штрафов, пени, неустоек за вычетом регуляторных резервов, скорректированный на стоимость рублевого акционерного капитала, норматив достаточности капитала, риск-вес активов и количество дней в году.

$$\text{ЧЗК} = ((\text{ОСЗ}_i + \text{Проценты}_i + \text{ШПН}_i) - \text{Регуляторные резервы}_i) \times \frac{\text{CoE}}{365} \times H1 \times RWA_i$$

, где ОСЗ - остаток ссудной задолженности, ШПН - штрафы, пени и неустойки, CoE - стоимость капитала, H1 - норматив достаточности собственного капитала (H1.0).

Изменение компонентов объема задолженности и RWA приводят к изменению Чистых затрат на капитал.

Кроме того, значение риск-веса участвует также в расчете капитала:

$$\text{Собственный капитал}_i = ((\text{ОСЗ}_i + \text{Проценты}_i + \text{ШПН}_i) - \text{Регуляторные резервы}_i) \times H1 \times RWA_i$$

А затраты на фондирование, в свою очередь, зависят от объема капитала Банка:

$$\text{Хондирование}_i = ((\text{ОСЗ}_i + \text{Проценты}_i + \text{ШПН}_i) - \text{Регуляторные резервы}_i - \text{Собственный капитал}_i) \times \frac{\text{ETC}}{365}$$

, где ETC - единая трансфертная ставка.

Финансовый эффект от реализации модели = изменение веса и значений компонент стоимости капитала по процессу урегулирования проблемной задолженности

В завершении темы про Резервы и Капитал, нужно отметить, что мониторинг подобных эффектов крайне важен. Необходимо контролировать изменение драйверов эффекта (например, стабильности уровня сбора задолженности, расходов на сбор задолженности, доля кредитов, уходящих в дефолт) и в соответствии с изменениями корректировать расчет эффекта. Также стоит обращать внимание на эффекты второго порядка, так как существует тесная связь между показателями, участвующими в расчете финансового результата.

Набор метрик для оценки финансового эффекта в зависимости от целей, задач

Набор задач и метрик расширяется по мере роста зрелости ИИ-оценки: от окупаемости отдельной инициативы (АФИИНА I) — к портфельному управлению (АФИИНА II) и системной оценке (АФИИНА III). Унифицированные (универсальные) метрики задаются здесь, на уровне; отраслевая специфика вынесена в отраслевые разрезы внутри соответствующего уровня и не дублирует универсальные метрики.

АФИИНА I

Оценка отдельной ИИ-инициативы. Задача — внешняя отчетность и подсчет окупаемости.

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Внешний репортинг	PnL отчётного года	Влияние инициативы на прибыль и убытки в рамках отчётного периода	Когда требуется верифицируемость и сопоставимость с другими статьями отчётности
Базовый подсчёт окупаемости	PnL 12 месяцев вперёд	Прогнозируемый эффект на прибыль в следующие 12 месяцев	Короткий горизонт планирования; рынок волатильный; инициатива не требует долгосрочного прогноза
Эффект через ресурсную экономию	Виртуальный PnL	Оценка через сэкономленные ресурсы, не отражённые напрямую в P&L	Когда эффект реализуется через ненайм, высвобождение мощностей или оптимизацию без прямого попадания в P&L

О метриках уровня 1:

PnL отчётного года — наиболее строгая метрика с точки зрения внешней отчётности. Она требует, чтобы эффект был реализован и зафиксирован в рамках одного отчётного периода, что делает её удобной для сопоставления с другими статьями P&L, но неподходящей для инициатив с отложенным эффектом.

PnL 12 месяцев отличается от отчётного года тем, что «окно» оценки скользящее и не привязано к календарному году. Это позволяет запустить оценку в любой момент и получить сопоставимый горизонт для всех инициатив портфеля, независимо от даты их старта. Метрика проста в расчёте и понятна всем участникам процесса — от бизнес-юнита до финансового контролёра.

Виртуальный PnL фиксирует эффекты, которые не отражаются напрямую в отчёте о прибылях и убытках. Типичный пример — ненайм: если ИИ позволяет не нанимать 10 сотрудников, эта экономия не появится в P&L как статья дохода, но является реальным финансовым эффектом. Важно применять виртуальный PnL осторожно: он не подходит для внешнего репортинга и требует чёткого методологического обоснования, иначе легко зависеть оцениваемый эффект.

Ключевое правило: для внешнего репортинга использовать только метрики в терминах PnL — они обеспечивают сопоставимость и верифицируемость. CLTV, NPV и виртуальный PnL допустимы исключительно для внутренней оценки.

АФИИНА II

Оценка портфеля ИИ-инициатив. Задача — приоритизация, повышение конкурентоспособности, оценка косвенных эффектов.

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Приоритизация долгосрочных или капиталоемких инициатив	NPV	Чистая приведённая стоимость — дисконтированный многолетний поток эффектов	Жизненный цикл инициативы более 1–2 лет; важно сравнить инициативы с разными временными профилями эффекта
Сравнение инициатив по отдаче на вложения	ROI	Отношение совокупного эффекта к инвестициям	Ограниченный бюджет; нужен единый критерий отбора внутри портфеля
Инициативы, влияющие на клиентскую ценность	CLTV	Изменение пожизненной ценности клиента под влиянием ИИ	Только для внутренней оценки; компания располагает корректно рассчитанной метрикой LTV
Управление рисками (фрод, дефекты, просрочка)	Loss prevention	Объём предотвращённых потерь	Когда прямой прирост дохода отсутствует, а эффект выражается в снижении убытков
Инициативы, влияющие на клиентский опыт	CSI / NPS / CSAT	Клиентская удовлетворённость и лояльность	Используется в связке с финансовыми метриками — не замещает их, а обосновывает причинно-следственную связь
Автоматизация и ускорение процессов	Time-to-XXX	Скорость достижения операционного результата (решение, доставка, ответ)	Когда эффект — в ускорении; переводится в PnL через экономию OPEX или рост пропускной способности

О метриках уровня 2:

NPV позволяет корректно сравнивать инициативы, которые начинают приносить эффект в разное время и с разным профилем. Например, инициатива А даёт быстрый, но небольшой эффект в первый год, а инициатива Б — значительно больший эффект, но начиная со второго-третьего года. PnL 12 месяцев покажет преимущество А, тогда как NPV может показать преимущество Б. Ставка дисконтирования отражает стоимость капитала и риски компании.

CLTV (Customer Lifetime Value) — метрика, отражающая совокупную ценность клиента за всё время его взаимодействия с компанией. Изменение CLTV под влиянием ИИ-инициативы показывает не только текущий эффект, но и долгосрочное влияние на клиентскую базу. Важно: CLTV не является метрикой для внешней отчётности, поскольку её расчёт существенно варьируется между компаниями и не верифицируется внешними аудиторами.

CSI, NPS и CSAT фиксируют изменение клиентского опыта, которое не всегда сразу трансформируется в финансовый результат. Эти метрики используются как «промежуточные» — они показывают, что инициатива движется в правильном направлении, пока полный финансовый эффект ещё не накопился. Для корректной интерпретации рекомендуется построить драйверную модель: насколько рост NPS на X пунктов исторически коррелирует с изменением Retention и, соответственно, PnL.

Time-to-XXX — семейство метрик, измеряющих скорость ключевых процессов: Time-to-Decision, Time-to-Market, Time-to-Response и т.д. Сам по себе Time-to-XXX является операционной метрикой, однако он конвертируется в финансовый эффект через снижение трудозатрат (OPEX) или увеличение пропускной способности процесса при той же ресурсной базе.

Приоритизация метрики для портфельной оценки

При необходимости сравнить инициативы по единому критерию рекомендуется выбрать **одну** приоритетную метрику на основе следующей логики:

- **PnL 12 мес.** — рынок волатильный, горизонт планирования короткий, инициатива краткосрочная. Метрика проста в расчёте и верификации.
- **NPV** — инициатива капиталоемкая или с жизненным циклом более 1–2 лет. Позволяет корректно сравнивать инициативы с разными временными профилями эффекта.
- **CLTV** — инициатива направлена на клиентскую ценность и компания располагает корректно рассчитанной метрикой LTV. Только для внутренней приоритизации.
- **Виртуальный PnL** — эффект реализуется через ненайм или ресурсную оптимизацию, не отражаемую напрямую в P&L.

Использование нескольких метрик одновременно для сравнения инициатив между собой затрудняет приоритизацию и нередко приводит к противоречивым выводам. Если для конкретной инициативы важны несколько измерений эффекта — рекомендуется зафиксировать **одну** как основную для ранжирования, остальные — как дополнительные для контроля.

Отраслевые метрики по отраслям

Ниже — специфические отраслевые метрики, применяемые на уровне портфельной и стратегической оценки. Универсальные метрики (NPV, ROI, CLTV, PnL) не повторяются — они заданы выше на уровне. Все метрики переводятся в денежное выражение через документированную цепочку драйверов.

III.1. Финансовая индустрия

задача	метрика	описание	когда использовать
Риск-взвешенная доходность капитала	eROE / RAROC	Ожидаемая чистая прибыль, отнесённая к потреблению капитала: $eROE = \frac{\text{ожидаемая чистая прибыль}}{(RWA \times CAR)}$, где RWA — активы, взвешенные по риску, CAR — норматив достаточности капитала	Когда ИИ-модель (скоринг, риск-модели, модель ПДН) снижает потребление капитала при той же прибыли; эффект оценивается как $\Delta eROE$
Стоимость кредитного риска	Cost of Risk / эффект на резервы	Отчисления в резервы (РВПС) на портфель; улучшение PD/LGD-моделей снижает резервы, а также ускоряет и увеличивает сбор просроченной задолженности	Когда ИИ улучшает качество скоринга и взыскания; разграничивать временный и постоянный эффект, оценивать life-time (принцип неттирования)
Высвобождение ликвидности	Доход от высвобождённой ликвидности	Объём ликвидности, высвобождённой ИИ (норматив ликвидности, СДО в кассовых центрах и банкоматах), умноженный на ставку размещения (Overnight)	Когда ИИ оптимизирует остатки и нормативы ликвидности; ставка экзогенна, при отрицательной ставке эффект отрицателен

- **eROE / RAROC** — ключевая метрика риск-взвешенной эффективности: ИИ создаёт стоимость не только ростом прибыли, но и снижением потребляемого капитала. Пример из методики (модель ПДН, Альфа-Банк, премия GIA 2024): eROE растёт с 35% до 58%, $\Delta eROE \approx 23$ п.п. за счёт снижения RWA.
- **Cost of Risk** — эффект на резервы разграничивать по знаку и горизонту (временный / постоянный) и оценивать life-time с учётом принципа неттирования; эффект не признаётся при изменении учётной политики банка или норм регулятора.
- **Высвобождение ликвидности** — анализировать ошибки первого и второго рода: ошибка в оценке свободной ликвидности может привести к дефициту и привлечению средств по более высокой ставке.

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Оценка масштаба эффекта на платформе	GMV (Gross Merchandise Value)	Суммарная стоимость всех товаров и услуг, реализованных через платформу за период	Используется как верхнеуровневый ориентир при внешнем репортинге; для атрибуции эффекта конкретной ИИ-инициативы требует декомпозиции на составляющие драйверы — частоту заказов, средний чек, конверсию
Оценка экономики единицы транзакции	Contribution Profit	Маржинальная прибыль с одного заказа после вычета прямых переменных затрат: логистика, выплаты партнёрам, обработка	Приоритетная метрика оценки ИИ-инициатив на уровне транзакции; объединяет эффекты роста выручки и снижения затрат в единый показатель; используется при защите инициатив перед финансовым блоком
Оценка производительности операционного ресурса	Order Per Hour (OPH)	Количество заказов, выполненных операционной единицей (курьером, диспетчером) за один час рабочего времени	Когда ИИ оптимизирует маршрутизацию, диспетчеризацию или планирование смен; позволяет оцифровать операционный эффект без сведения к ПШЕ
Оценка удельных операционных затрат	Cost Per Order (CPO)	Прямые операционные затраты на выполнение одного заказа	Ключевая метрика для оценки ИИ-маршрутизации; снижение CPO при стабильном OPH — прямой OPEX-эффект. Состав затрат фиксируется до запуска и не меняется в ходе оценки
Оценка доходности ресурса	Money Per Hour	Выручка платформы или фактический заработок партнёра-курьера за один час рабочего времени	Применяется в паре с OPH: рост OPH без роста Money Per Hour означает перераспределение нагрузки без создания дополнительной ценности
Оценка качества исполнения заказов	Уровень опозданий (Late Rate)	Доля заказов, доставленных позже обещанного клиенту времени	Когда ИИ применяется в прогнозировании времени доставки или управлении пиковой нагрузкой; служит прокси-метрикой для NPS и долгосрочного оттока; переводится в деньги через когортный анализ возвратности
Оценка эффекта динамического ценообразования	Surge Fee	Надбавка к базовой цене заказа, применяемая ИИ в периоды пиковой нагрузки для балансировки спроса и предложения	Оценивается двусторонне: рост выручки с заказа сопоставляется с изменением конверсии и общего числа размещённых заказов
Оценка монетизации платформы	Service Fee / Take Rate	Комиссия платформы с каждого заказа; доля выручки платформы в суммарном GMV	Когда ИИ влияет на масштаб обработанных заказов или условия работы с партнёрами
Оценка эффективности воронки продаж	Conversion Rate (CR) по этапам	Доля пользователей, совершивших целевое действие на каждом шаге: открытие → поиск → корзина → оплата	Когда ИИ применяется в персонализации и ранжировании; оценивается по каждому переходу отдельно — ИИ может улучшить один этап и ухудшить другой

- **GMV** для изолированной оценки требует декомпозиции: рост GMV одновременно определяется маркетинговыми кампаниями, ассортиментной политикой и сезонностью. Инкрементальный вклад измеряется через A/B тест, свитчбэк-эксперимент или глобальный холдаут.
- **Contribution Profit** — наиболее репрезентативная метрика для оценки ИИ в e-com; состав затрат фиксируется до запуска инициативы и не меняется в ходе оценки.
- **Order Per Hour** нормируется на тип заказа и географическую зону: рост OPH при одновременном снижении CPO — наиболее убедительное подтверждение операционного ИИ-эффекта.
- **Late Rate** переводится в деньги через когортный анализ: разница в возвратности между пользователями, получившими опоздание, и контрольной группой, умноженная на средний Contribution Profit заказа.
- **Surge Fee** оценивается двусторонне — оптимум максимизирует Contribution Profit, а не выручку с заказа.

Методологическое замечание. На платформах с высокой частотой А/Б тестирования совокупный эффект фичей, измеренных поочерёдно в изолированных экспериментах, систематически расходится с реальным инкрементальным улучшением: фичи взаимодействуют между собой. Для корректной оценки рекомендуется поддерживать постоянную контрольную группу (глобальный холдаут), не получающую ни одного ИИ-улучшения.

II.III. Транспорт и логистика

Для компаний транспортного сектора характерна зрелая система инвестиционной оценки: все ИИ-инициативы проходят через расчёт NPV, IRR, дисконтированного срока окупаемости и индекса доходности с учётом WACC.

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Оценка коммерческой загрузки	Passenger Load Factor (PLF)	Доля занятых кресел от общего предложения мест на выполненных рейсах	Когда ИИ применяется в revenue management; оценивается в паре с Yield — высокий PLF при низком Yield означает заполнение дешёвыми билетами
Оценка доходности провозной ёмкости	RASK (Revenue per Available Seat Km)	Выручка в расчёте на одно доступное кресло-километр	Итоговая метрика ИИ в revenue management; рост RASK при стабильном CASK — прямой PnL-эффект
Оценка затрат на провозную ёмкость	CASK (Cost per Available Seat Km)	Совокупные операционные затраты в расчёте на одно доступное кресло-километр	Снижение CASK — целевой эффект ИИ в маршрутизации, управлении ТО и наземных операциях
Оценка качества исполнения рейсов	On-Time Performance (OTP)	Доля рейсов, выполненных по расписанию с отклонением не более 15 минут	Задержки влекут компенсационные выплаты, штрафы за слоты и сверхурочные расходы; каждая задержка имеет измеримую финансовую стоимость
Оценка использования флота	Коэффициент использования ВС (Utilization Rate)	Среднее время налёта воздушного судна в часах в сутки	Рост утилизации без деградации OTP — реальное операционное улучшение
Оценка топливной эффективности	Удельный расход топлива (Fuel Burn per ASK)	Расход авиационного топлива в расчёте на одно доступное кресло-километр	Топливо составляет 30–40% CASK в российской авиации — приоритетная статья OPEX-эффектов
Оценка затрат на техническое обслуживание	MRO Cost per Flight Hour	Затраты на ТО и ремонт в расчёте на один час налёта	При оценке разграничивать плановые и внеплановые остановки: ИИ воздействует прежде всего на внеплановые

- **PLF и Yield** анализируются совместно: рост PLF за счёт снижения цен не является положительным ИИ-эффектом, поэтому корректная оценка строится на паре $PLF \times Yield \rightarrow RASK$.
- **OTP** имеет прямой финансовый выход — для перевода в деньги рассчитывается средняя стоимость задержки на рейс по историческим данным.
- **MRO Cost per Flight Hour**: внеплановые остановки значительно дороже плановых, поэтому ведётся отдельный учёт плановых и внеплановых затрат.
- **PI (Profitability Index)** ценен при портфельной приоритизации: нормируя эффект на объём ресурсов, позволяет корректно сравнивать крупные инфраструктурные проекты с точечными улучшениями.

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Оценка доходности абонентской базы	ARPU (B2C и B2B отдельно)	Средняя ежемесячная выручка на одного активного абонента	Сегменты анализируются строго отдельно — ARPU B2B кратно выше и определяется иными драйверами
Оценка удержания клиентов	Churn Rate	Доля абонентов, прекративших пользование услугами за период	Перевод в деньги: $\Delta \text{Churn} (\%) \times \text{количество абонентов} \times \text{ARPU} \times \text{средний ожидаемый срок жизни абонента}$
Оценка эффективности привлечения	CAC (Customer Acquisition Cost)	Совокупные затраты на маркетинг и продажи на одного нового абонента	Снижение CAC отслеживается совместно с Churn Rate когорты — ухудшение профиля нивелирует экономию
Оценка долгосрочной ценности клиента	CLTV (Customer Lifetime Value)	Суммарный дисконтированный доход абонента за весь срок пребывания	Ориентир качества портфеля: $\text{CLTV} / \text{CAC} \geq 3$
Оценка клиентского сервиса	FCR (First Contact Resolution)	Доля обращений, разрешённых при первом контакте без повторных обращений	Рост FCR напрямую снижает OPEX контактного центра и позволяет обосновать ненайм при росте входящего потока
Оценка сетевой эффективности	Network Utilization	Доля использованной пропускной способности в часы пиковой нагрузки	Повышение утилизации без деградации QoS = обслуживание большего трафика без CAPEX на расширение

Churn Rate — приоритетный объект воздействия ИИ: стоимость привлечения нового абонента в 5–7 раз превышает стоимость удержания.

CAC отслеживается совместно с профилем когорты: снижение CAC при ухудшении Churn и ARPU может давать отрицательный совокупный эффект.

FCR — прямой финансовый выход: снижение повторных обращений $\times$ средняя стоимость обращения = OPEX-эффект.

Network Utilization — финансовый эффект выражается как отложенный CAPEX: объём инвестиций в расширение сети, которые не потребуются благодаря ИИ-оптимизации.

Методологическое замечание. Телекоммуникационные компании с присутствием в нескольких сегментах сталкиваются с принципиально разными метриками в каждом направлении. Рекомендуется разработать единый стандарт оценки на уровне финансового директора с адаптацией под каждый сегмент — вместо независимых методологий на уровне бизнес-подразделений.

II.V. АПК, энергетика и промышленность

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Оценка надёжности оборудования	КТГ (коэффициент технической готовности)	Доля времени, в течение которого единица оборудования находится в работоспособном состоянии	Когда ИИ применяется для предиктивного ТО; каскад: $\Delta \text{КТГ} \rightarrow \text{доп. часы работы} \rightarrow \text{объём выпуска (тонн)} \rightarrow \text{финансовый эффект (тонны} \times \text{маржа руб./т)}$
Комплексная оценка производственной эффективности	OEE (Overall Equipment Effectiveness)	Интегральный показатель: Готовность $\times$ Производительность $\times$ Качество продукции	Предпочтительнее КТГ, когда ИИ воздействует одновременно на несколько аспектов; декомпозиция позволяет точно атрибутировать канал создания ценности
Оценка выхода ценного компонента	Выход металла / Recovery Rate	Доля полезного компонента, извлечённого из сырья при обогащении или металлургическом переделе	Для горно-металлургических производств; нормировать на содержание металла во входящем сырье — вариативность рудной базы маскирует реальный ИИ-эффект
Оценка конкурентоспособности по затратам	C1 Cost (прямая себестоимость на тонну)	Прямые операционные затраты на производство тонны продукции без амортизации, налогов и процентных расходов	Отраслевой стандарт для сопоставления с глобальными аналогами; снижение C1 Cost улучшает позицию на мировой кривой затрат
Оценка энергоэффективности	Удельный расход электроэнергии (кВт·ч/т)	Потребление электроэнергии на единицу произведённой продукции	Когда ИИ оптимизирует энергоёмкие процессы; приоритетна при высокой доле энергозатрат в себестоимости
Оценка ценности продуктового качества	Grade Mix / доля продукции высшего сорта	Доля продукции, соответствующей premium-характеристикам по чистоте, составу или стандартам	Разница в цене между сортами — прямой доходный эффект без роста объёма выпуска
Оценка интенсивности добычи (нефтегаз)	Дебит скважины (т/сут или тыс. м³/сут)	Объём нефти или газа, добытый из скважины в единицу времени; в т/сут для нефти и тыс. м³/сут для газа	Когда ИИ применяется для оптимизации режима эксплуатации скважин (газифит, ГРП, управление штуцером); каскад: $\Delta \text{дебит (т/сут)} \times \text{дни} \times \text{цена реализации} = \text{финансовый эффект}$
Оценка удельных затрат на добычу (нефтегаз)	Lifting Cost (руб./т н.э. или долл./барр.)	Прямые операционные затраты на подъём одной тонны нефтяного эквивалента; не включает геологоразведку, налоги и амортизацию	Отличается от C1 Cost меньшим составом затрат — только операционные расходы добычи; снижение Lifting Cost через ИИ-оптимизацию режимов — прямой OPEX-эффект, особенно значим при низких ценах на нефть
Оценка эффективности растениеводства (агро)	Урожайность (т/га)	Объём продукции, собранной с одного гектара посевной площади для конкретной культуры	Когда ИИ применяется в точном земледелии (переменное внесение удобрений, прогнозирование болезней, оптимизация сроков сева и уборки); каскад: $\Delta \text{урожайность (т/га)} \times \text{площадь (га)} \times \text{цена реализации (руб./т)} = \text{финансовый эффект}$
Оценка операционных затрат в растениеводстве (агро)	Себестоимость производства (руб./т по культуре)	Совокупные прямые затраты на производство одной тонны продукции: семена, удобрения, СЗР, механизация, труд	Когда ИИ снижает нормы внесения удобрений и СЗР или оптимизирует использование техники; снижение себестоимости при сохранении или росте урожайности — комплексный операционный эффект

- **КТГ** — разграничивать плановые и внеплановые простои: ИИ-диагностика воздействует прежде всего на внеплановые остановки, которые в 3–5 раз дороже плановых.
- **OEE** учитывает не только простои, но и скорость работы и долю брака; декомпозиция позволяет точно атрибутировать, через какой канал ИИ создаёт ценность.
- **Recovery Rate** нормируется на содержание металла во входящем сырье: вариативность рудной базы может маскировать реальное улучшение.

- **C1 Cost** для экспортно-ориентированных производителей оценивается и в рублях, и в долларовом эквиваленте — реальная конкурентоспособность определяется долларовой себестоимостью.
- **Дебит скважины** оценивается с учётом стадии разработки месторождения: на зрелых месторождениях базовый дебит снижается естественным образом, поэтому фактический дебит сравнивают с базовой кривой снижения (decline curve) по методу аналогов.
- **Lifting Cost**, в отличие от C1 Cost, охватывает только затраты непосредственно добычи (без геологоразведки, налогов на добычу и амортизации) — наиболее чувствительный индикатор для оценки ИИ в управлении режимом эксплуатации скважин.
- **Урожайность (т/га)** — необходимо контролировать влияние погоды: межгодовая вариативность из-за климата существенно превышает типичный ИИ-эффект. Рекомендуется сравнение с контрольными участками (рандомизированный полевой эксперимент) или нормирование на агрометеорологический индекс сезона.
- **Себестоимость (руб./т)** в агро требует разграничения переменных и постоянных затрат: ИИ снижает прежде всего переменные (химия, механизация), а постоянные (аренда, амортизация) в инкремент не включаются.
- **Документированный каскад.** Прямой экономический эффект раскрывается цепочкой, например: «ИИ повышает КТГ на 1,5 п.п. → +200 часов работы оборудования → при производительности X т/час и марже Y руб./т → Z млн руб.».
- **ФТЕ.** Фактическое сокращение подтверждается изменением штатного расписания; ненайм требует нормативной модели численности.
- **PI > 2** — жёсткий стандарт для быстрых операционных инициатив; для долгосрочных капиталоемких проектов применяется NPV с горизонтом 3–5 лет.

Методологическое замечание — эффект «непрожитой реальности». Когда ИИ ускоряет этапы инвестиционного проекта (проектирование, согласование, строительный контроль), контрфактическая реальность не была реализована, что затрудняет прямую атрибуцию. Рекомендуемый подход: (1) зафиксировать детальный базовый план до внедрения ИИ; (2) вести мониторинг отклонений с документированием причин ускорения; (3) атрибутировать ИИ-эффект только к этапам критического пути; (4) применять метод альтернативного моделирования. Без выполнения этих условий эффект остаётся гипотетическим.

Как измерять эффект, где нет денежных метрик? Косвенные стратегические метрики.

Блок про косвенные стратегические метрики относится ко второму уровню фреймворка АФИИНЫ, потому что именно на этом уровне компания выходит за рамки оценки одной инициативы и начинает работать с портфелем, где часть инициатив влияет на клиентский опыт, конкурентоспособность и операционную эффективность, не имея при этом прямого денежного выражения. На уровне 1 достаточно PnL: эффект прямой и верифицируемый. На уровне 2 появляются инициативы, где путь к финансовому результату опосредован — через NPS, Retention, Time-to-XXX — и именно здесь возникает необходимость в драйверной модели как инструменте перевода косвенного эффекта в финансовый.

Итак, ряд инициатив влияет на метрики, которые не имеют прямого денежного выражения: NPS, CSI, Retention, удовлетворённость клиентов и другие. При этом Retention всё же поддаётся переводу в PnL, поскольку напрямую влияет на доходы компании. Оставшиеся метрики — сложнее. Тем не менее это не означает, что такие инициативы нельзя оценить финансово. Метод — построить **драйверную модель**: связать косвенную метрику с финансовым результатом через цепочку промежуточных зависимостей.

Прокси-метрики могут быть переведены в PnL текущего периода с помощью исследовательских А/Б-тестов направленных на замер размена прокси-метрик с деньгами.

Пример:

Метрика лояльности – вовлеченность клиентов в экосистему

- 1 С помощью А/Б теста измеряем курс размена пункта лояльности с денежными метриками и получаем курс размена, например, Y.. Например, через применение прямых денежных стимулов: доп. кэшбэки/скидки (или наоборот уменьшение скидок и кэшбэков);
- 2 Через А/Б тест измеряем, как новый ИИ-продукт влияет на лояльность и PnL, получаем некий курс лояльности X. Важно помнить, что при этом курс размена не будет являться универсальным и нельзя транслировать его на другие кейсы;
- 3 Сравниваем курс лояльности X в ИИ-продукте с курсом размена Y (из п.1).

Если с помощью ИИ-продукта мы можем «купить» лояльность дешевле, чем через тривиальные изменения денежных стимулов, то ИИ положительно влияет на PnL.

Но есть инициативы, где крайне сложно, а порой и невозможно перевести метрику в финансовый эффект, но это не значит, что такие метрики не стоит учитывать при выборе ИИ проекта.

Мы предлагаем 3 группы потенциальных стратегических эффектов от ИИ с примерами метрик:

- 1 **Удовлетворенность / репутация**
 1. Рост лояльности
 2. % персонализации коммуникаций (от всех коммуникаций)
 3. Стоимость бренда / медиа охват с учетом сентимента
- 2 **Эффективность**
 1. Объем бизнеса на 1 сотрудника (например, доходы с КСБ на 1 сотрудника КИБ КСБ или количество закрытых обращений в КЦ на 1 сотрудника КЦ)
 2. % автономизации критических процессов (например, доля выданных кредитов через ИИ по сегментам клиентов)
 3. Охват ролей / сотрудников ИИ-помощниками
- 3 **Безопасность**
 1. Loss prevention
 2. % детекций фрода
 3. Объем утечек данных (количество утечек X количество затронутых клиентов)

Упомянутые метрики необходимо рассматривать индивидуально под каждый конкретный проект и, возможно, адаптировать их с учетом специфики и целей. Эти показатели не являются универсальными или обязательными для использования, но они основаны на многолетнем опыте оценки проектов. Поэтому подход к выбору и применению метрик должен быть гибким и ориентированным на достижение максимальной ценности для бизнеса в каждом отдельном случае.

АФИИНА III

Оценка влияния ИИ на устойчивость компании и отрасли. Задача — стратегическое планирование, ИИ-стратегия, отраслевой бенчмаркинг.

Универсальные метрики (применимы к большинству секторов)

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Долгосрочный P&L-эффект ИИ-трансформации	Накопленный PnL (3–5+ лет)	Совокупный эффект на прибыль за стратегический горизонт с учётом деградации моделей	Оценка стратегической ценности ИИ-программы в целом; обоснование крупных инвестиций
Операционная надёжность и устойчивость	Снижение простоев / MTBF	Снижение простоев или рост среднего времени между отказами	Производство, инфраструктура, логистика — инициативы предиктивного обслуживания
Качество процессов и продуктов	Доля дефектов / доля ошибок	Доля дефектов или ошибок в процессах	Промышленность, здравоохранение, логистика; снижение rate переводится в Loss prevention

Отраслевые метрики по отраслям

На системном уровне отраслевые метрики носят долгосрочный, регуляторный или капитальный характер и оцениваются на горизонте нескольких лет; нередко они выступают жёсткими ограничениями реализуемости инициативы, а не только источником эффекта.

III.I. Финансовая индустрия

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Кредитные риски и модели оценки	Влияние на резервы (IFRS 9 / РСБУ)	Изменение объёма резервов под кредитные убытки за счёт улучшения PD/LGD-моделей	Банки и финансовые организации; оценка эффекта скоринговых ИИ-моделей
Регуляторная эффективность	Влияние на капитал и RWA	Снижение активов, взвешенных с учётом риска; высвобождает регуляторный капитал	Системно значимые банки; инициативы по оптимизации кредитного портфеля

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Вклад сектора в ценовую эффективность экономики	Розничная наценка сектора / вклад в потребительские цены	Средняя торговая наценка распределительного слоя и её вклад в конечную цену товаров	Когда массовый ИИ (ценообразование, прогноз спроса, оптимизация цепочки) снижает издержки и неэффективность розницы — снижается «розничная наценка» на товары во всей экономике (ритейл-аналог «логистического налога»)
Ресурсная эффективность и продовольственная безопасность	Товарные потери и списания в отрасли (waste rate)	Доля потерь, списаний и излишков товара на уровне сектора, включая скоропорт	Когда массовый ИИ-прогноз спроса и управление запасами сокращают потери и перепроизводство в масштабе отрасли — системная ресурсная эффективность
Структурная капиталоемкость сектора	Оборотный капитал в запасах (доля от оборота отрасли)	Объём капитала, замороженного в товарных запасах сектора, относительно его оборота	Когда массовый ИИ позволяет обслуживать тот же оборот меньшими запасами — высвобождается оборотный капитал во всём секторе, растёт отдача на капитал
Структурный сдвиг канала	Проникновение e-com (доля цифрового канала в обороте розницы)	Доля онлайн-канала в общем розничном обороте	Когда ИИ (персонализация, логистика, ценообразование) ускоряет переход розницы в онлайн — системный индикатор цифровой трансформации отрасли

- **Розничная наценка / ценовая эффективность** — флагманская системная метрика: розница как распределительный слой между производством и потреблением, поэтому её удешевление за счёт ИИ снижает конечные цены во всей экономике (аналог «логистического налога» в транспорте).
- **Товарные потери** — на системном уровне это ресурсная эффективность и вклад в продовольственную безопасность, а не операционная метрика отдельной сети.
- **Оборотный капитал в запасах** — отличать от оборачиваемости запасов отдельной компании (уровень II): здесь речь о капитале, замороженном в запасах всей отрасли, как структурной характеристике сектора.

III.III. Транспорт и логистика

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Конкурентоспособность отрасли по издержкам	Удельные логистические издержки (себестоимость тонно-километра; доля логистики в выручке / в ВВП)	Совокупная стоимость логистики на единицу перевозки на уровне сети или отрасли	Когда массовое внедрение ИИ (маршрутизация, консолидация грузов, автономный транспорт) снижает стоимость логистики во всей экономике; удешевление повышает ценовую конкурентоспособность отрасли и производимых товаров
Эффективность использования активов отрасли	Утилизация провозных мощностей / доля порожнего пробега	Загрузка флота или парка и доля пробега без груза на уровне сети	Отрасль перевозит тот же объём меньшим парком — падает капиталоемкость, растёт отдача на активы (ROIC)
Системная надёжность транспортной сети	Аварийность / надёжность доставки в срок на уровне сети	Частота инцидентов и доля своевременных доставок по транспортной системе в целом	ИИ-предиктивное обслуживание и управление трафиком повышают устойчивость цепочек поставок в экономике

- **Удельные логистические издержки** — ключевая метрика конкурентоспособности: логистика входит в себестоимость почти всех товаров, поэтому её системное удешевление за счёт ИИ снижает «логистический налог» на экономику и повышает конкурентоспособность отрасли в целом.
- **Утилизация провозных мощностей / порожний пробег** — при массовой ИИ-оптимизации отрасль перевозит тот же объём меньшим парком: снижается капиталоемкость и растёт отдача на вложенный капитал.
- **Системная надёжность сети** — предиктивное обслуживание и интеллектуальное управление трафиком снижают сбои на уровне всей транспортной системы, повышая устойчивость цепочек поставок.

III.IV. ИКТ

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Капиталоемкость отрасли	Capex-to-Sales / capex на единицу трафика	Отношение капитальных затрат к выручке и капзатраты на единицу переданного трафика на уровне сети	Когда ИИ-автоматизация планирования и эксплуатации сети (self-optimizing networks, AI-RAN) снижает капзатраты на строительство и поддержание сети; снижение капиталоемкости — структурное конкурентное преимущество отрасли
Удельная стоимость передачи трафика	Cost per GB (себестоимость передачи единицы трафика)	Совокупная стоимость передачи одного гигабайта (или бита) на уровне сети	Телеком-аналог удельных логистических издержек: ИИ удерживает падение себестоимости передачи данных при кратном росте трафика — мера способности отрасли обслуживать цифровую экономику
Системная надёжность и устойчивость сети	Availability / QoS на уровне сети	Доступность и качество услуг по сети в целом (а не отдельные B2B-SLA)	Устойчивость национальной телеком-инфраструктуры; ИИ-предиктивное обслуживание и управление трафиком снижают системные сбои

- **Капиталоёмкость сети** — телеком крайне капиталоемок; ИИ-автоматизация сети снижает сарех на её строительство и эксплуатацию, что структурно повышает отдачу на капитал в отрасли.
- **Удельная стоимость передачи трафика** — при взрывном росте трафика именно удержание падения себестоимости передачи данных определяет конкурентоспособность отрасли и её способность обслуживать цифровую экономику.
- **Системная надёжность сети** — на уровне отрасли QoS/доступность отражает устойчивость инфраструктуры; ИИ-предиктивное обслуживание и интеллектуальное управление трафиком снижают системные сбои.

III.V. АПК, энергетика и промышленность

ЗАДАЧА	МЕТРИКА	ОПИСАНИЕ	КОГДА ИСПОЛЬЗОВАТЬ
Оценка эффективности извлечения запасов (нефтегаз)	КИН (коэффициент извлечения нефти / Oil Recovery Factor)	Доля извлечённых запасов от геологических запасов месторождения на текущий момент	Когда ИИ применяется в управлении пластовым давлением и разработке месторождений; прирост КИН на 1 п.п. при масштабе месторождения эквивалентен миллиардам рублей дополнительной добычи в горизонте нескольких лет
Оценка промышленной безопасности	LTIFR (Lost Time Injury Frequency Rate)	Число производственных травм с потерей рабочего времени на 1 млн человеко-часов	Финансовый выход: страховые выплаты, регуляторные штрафы, потери от вынужденных остановок производства
Оценка экологических и регуляторных рисков	Удельные выбросы загрязняющих веществ (на тонну)	Объём выбросов ключевых загрязняющих веществ на единицу продукции	Финансовый выход: снижение экологических платежей, штрафов и затрат на природоохранные мероприятия

- **КИН** — стратегическая метрика: прирост даже на десятые доли процента на крупном месторождении создаёт значительный долгосрочный эффект; для атрибуции требуется гидродинамическое моделирование пласта с явным сравнением сценариев «с ИИ» и «без ИИ».
- **LTIFR и удельные выбросы** — регуляторные и ESG-метрики: финансовый выход через снижение страховых выплат, регуляторных штрафов и природоохранных платежей; на системном горизонте влияют на стоимость капитала и доступ к финансированию через ESG-рейтинг.

R&D

Инвестиции в R&D — не обязательный элемент ИИ-стратегии для каждой компании. На уровне 1 компании, как правило, не занимаются R&D: их задача — внедрить существующие решения и получить измеримый финансовый эффект в краткосрочном горизонте. Вопрос об инвестициях в R&D становится актуальным начиная с уровня 2 — когда компания переходит от управления отдельными инициативами к управлению портфелем и начинает думать о долгосрочной конкурентоспособности.

Прикладной R&D - АФИИНА II

«Прикладной» R&D — развитие бизнеса, требующее инвестиций на данный момент, которое потенциально принесёт денежный эффект через некоторое время. Имеет бизнес-заказчика. В рамках прикладного R&D можно отдельно выделить **платформенный ИИ**, который создаётся специально для владельцев платформ.

Прикладной R&D встраивается в логику среднесрочной стратегической оценки, потому что имеет бизнес-заказчика, измеримый горизонт (1–3 года) и конкретный ожидаемый эффект на метрики бизнеса. Именно на уровне 2 компания начинает управлять портфелем, а не отдельными инициативами, и прикладной R&D — полноправная часть этого портфеля.

Критерии прикладного R&D:

- Основная метрика — конверсия в прод
- Наличие бизнес-заказчика, техническая реализуемость, научная новизна
- Наличие прямого или косвенного бизнес-эффекта / влияния на метрики бизнеса, включая платформенные метрики (в т.ч. GR)
- Горизонт планирования 1–3 года

Как учитываем в эффекте от ИИ: показываем кумулятивно расходы в PnL по определённому R&D.

Как измерять эффект: ценность проекта оценивается в момент перехода проекта в продукт и получения финансового эффекта через его продажи. При этом у команды могут быть не только успешные проекты — эффект от успешных проектов должен «перекрывать» неудавшиеся запуски.

ВОЗМОЖНЫЕ КРІ

Процент конверсии в успешные проекты / доля успешных проектов

Количество продаж успешного продукта

Фундаментальный R&D - АФИИНА III

«Фундаментальный» R&D — развитие науки и поддержка научной активности, в т.ч. развитие Practical AGI. Фундаментальные исследования не имеют определённого бизнес-результата и не поддаются оценке через PnL или NPV. Их ценность реализуется в горизонте 10+ лет и влияет на технологическую устойчивость компании, способность конкурировать на рынке в долгосрочной перспективе и создавать принципиально новые продукты.

Критерии фундаментального R&D:

- Основная метрика — научная публикация
- Фундаментальные исследования
- Отсутствие определённого бизнес-результата по итогам исследования
- Горизонт планирования до 10 лет

Как учитываем в эффекте от ИИ: отдельная статья, расходы не включаем в эффект от ИИ.

Как измерять эффект: ценность проекта возможно оценить внешним сообществом учёных.

ВОЗМОЖНЫЕ КРІ

Количество и статус публикаций (A*)

Цитируемость

Инновационность (оценивается учёными)

Как распределить бюджет на R&D?

Как один из интересных примеров распределения бюджета на R&D, можно привести пример опыта компании **Huawei**. Huawei ежегодно инвестирует около 180 миллиардов юаней в исследования и разработки (при выручке 2024 года R&D составил около 20%), при этом важное разделение средств компании отражает разные цели и сроки исследований. Около 60 миллиардов юаней направляется на фундаментальные исследования — это долгосрочные проекты, результаты которых не поддаются немедленной оценке по критериям эффективности или коммерческой выгоды. Эти исследования создают базу для технологических прорывов и новых знаний, которые могут проявиться только через годы или даже десятилетия, но без них невозможен устойчивый инновационный рост. Еще 120 миллиардов юаней тратятся на прикладные исследования и разработку продуктов, где эффективность и результаты оцениваются по конкретным рыночным и коммерческим показателям. Это позволяет компании реализовывать новые технологии на практике, создавать востребованные продукты и сохранять конкурентоспособность на быстро меняющемся рынке. Такое четкое разделение и баланс между фундаментальным и прикладным R&D позволяют Huawei не только строить инновации с опорой на глубокие научные открытия, но и оперативно реагировать на требования промышленности и бизнеса, что особенно важно в условиях текущих глобальных вызовов.

Второй пример — Meta Platforms (2025). Если кейс Huawei показывает высокую R&D-интенсивность в производственной компании, то Meta — пример технологической компании, где разделение на прикладной и долгосрочный R&D видно прямо в сегментной отчетности. За 2025 год выручка Meta составила **200,97 млрд долларов**, а расходы на R&D — **57,37 млрд долларов**, то есть 28,5% выручки (рост на 31% год к году). Это выше, чем у Huawei (21,8% выручки): Meta направляет в исследования и разработку почти каждый третий доллар выручки.

Разделение по горизонту эффекта у Meta выражено через два бизнес-сегмента. Прикладной R&D сосредоточен в сегменте Family of Apps — это лента, рекомендательные системы, рекламные алгоритмы и основные продукты. У него есть прямой бизнес-заказчик и измеримый эффект: этот R&D обслуживает выручку в 201 млрд долларов и остаётся прибыльным, что соответствует логике уровня АФИИНА II (конверсия в продукт, горизонт 1–3 года). Долгосрочные и фундаментальные вложения сосредоточены в сегменте **Reality Labs** — направления AR/VR, метавселенной и «personal superintelligence» с горизонтом отдачи 10 и более лет. В 2025 году этот сегмент принёс лишь 2,21 млрд долларов выручки при **операционном убытке 19,19 млрд долларов**; накопленный убыток сегмента превысил 70 млрд долларов, а компания прямо закладывает сопоставимые убытки и на 2026 год. На этот долгосрочный контур приходится около 18% всех затрат компании. По логике методологии это уровень АФИИНА III: расходы не привязаны к ближайшему бизнес-результату и оцениваются как стратегическая ставка, а не как окупаемая в среднесрочном горизонте инициатива.

Для методологии этот пример важен двумя выводами. Во-первых, соотношение прикладного и долгосрочного R&D — это стратегический выбор компании, а не отраслевой норматив: у Huawei пропорция прикладного к фундаментальному составляет примерно 2:1, тогда как Meta направляет на долгосрочную ставку порядка трети R&D-бюджета. Во-вторых, долгосрочный R&D может годами оставаться глубоко убыточным при измеримо прибыльном прикладном ядре, и именно поэтому его оценивают отдельными метриками уровня АФИИНА III, а не через окупаемость ближайших периодов.

Выбор типа и объёма R&D — стратегическое решение, которое каждая компания принимает самостоятельно исходя из своих целей, горизонта планирования и конкурентного контекста. Не существует универсально правильного соотношения фундаментального и прикладного R&D: компания, стремящаяся к технологическому лидерству в долгосрочной перспективе, может сделать ставку на фундаментальные исследования; компания, фокусирующаяся на быстром росте и захвате рынка, — сосредоточиться исключительно на прикладном направлении.

При этом далеко не каждая компания готова инвестировать в R&D в принципе — и это тоже нормально. R&D требует готовности к неопределённости, длинному горизонту окупаемости и высокой доле неудач. Многие

компании осознанно выбирают путь применения уже существующих технологий и решений, не создавая собственных, — и получают финансовый эффект от ИИ именно на этом пути.

Таким образом, вопрос не в том, обязана ли компания инвестировать в R&D, а в том, какая стратегия соответствует её целям.

Бенчмаркинг

Для эффективного управления цифровой трансформацией и оценки успешности внедрения искусственного интеллекта компаниям необходимо не только измерять внутренние результаты, но и сравнивать их с лучшими практиками на рынке. Это позволяет понять, где организация находится относительно других игроков отрасли и какой имеет потенциал для улучшения.

Бенчмарки — это ключевой инструмент стратегического анализа. Они служат ориентиром для оценки динамики собственных показателей относительно рынка и лидеров отрасли. Если метрики компании соответствуют или превосходят бенчмарки, значит, она движется в ногу с передовыми игроками. Если нет — это сигнал о наличии возможностей для роста и оптимизации.

В рамках методологии оценки финансового эффекта от ИИ предлагается использовать комплексный набор метрик с рыночными бенчмарками, которые позволяют провести объективное сравнение по нескольким направлениям (финансовый эффект от ИИ, операционная эффективность, темпы роста и инвестиционная активность, уровень применения ИИ). Далее представлены универсальные бенчмарки, которые возможно использовать в межотраслевой специфике, а ниже можно найти отраслевые бенчмарки.

Универсальные бенчмарки

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:
 - 20% — доля финансового эффекта от ИИ в EBIT по данным консалтинговых исследований ([McKinsey, The state of AI in early 2024](#)).

Эти данные указывают на то, что представители индустрии уже получают значительную долю своих финансовых результатов за счет ИИ. Для компаний, стремящихся к лидерству, целесообразно стремиться к достижению и превышению этих показателей.

02 Операционная эффективность

- Предлагаемая метрика:
 - % оптимизированного OPEX за счет ИИ / ГенИИ
- Потенциальный бенчмарк:
 - ~10% — среднее сокращение операционных расходов (OPEX) за счет внедрения ИИ и генеративных технологий (ГенИИ), согласно рыночным исследованиям ([McKinsey, The state of AI in 2023](#)).

Оптимизация OPEX через ИИ — один из наиболее быстрых способов повышения рентабельности, особенно в секторах с высокой долей рутинных процессов.

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - Среднегодовой темп роста со сдвигом на уровень зрелости
- Потенциальный бенчмарк:
 - 50% — среднегодовой темп роста финансового эффекта от ИИ в Сбере (2020–2024), что свидетельствует о ускоряющемся ROI и масштабировании решений ([Сбер, результаты по МСФО за 2024; интервью А. Ведяхина, Ведомости](#)).

- Предлагаемая метрика:
 - ROI
- Потенциальный бенчмарк:
 - 16% — средний ROI в бизнесе от ИИ ([SAP, AI drives 31% ROI](#))

- Предлагаемая метрика:
 - Темп роста инвестиций
- Потенциальный бенчмарк:
 - x1,5 — планы Сбера по **росту инвестиций в ИИ** до 2029 года([интервью Г.О.Грефа](#)); 88% — средний прирост инвестиций в ИИ в Т-Банке за 2023–2025 годы (внутренний прогноз по росту инвестиций в ИИ до конца 2025г.).

Эти данные подтверждают, что лидеры рынка не только активно инвестируют в ИИ, но и добиваются значительной отдачи. Компаниям важно не просто увеличивать бюджеты, но и контролировать их эффективность.

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % процессов / продуктов / бизнес-функций с ИИ
- Потенциальный бенчмарк:
 - ~80% — **доля бизнес-процессов**, использующих ИИ в Сбере ([выступление А. Ведяхина / обзор TAdviser](#)). Это показывает высокую степень интеграции технологий во все аспекты деятельности.

Специфичные бенчмарки по группам отраслей экономики

Финансовая отрасль

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:

- 24% — **доля финансового эффекта от ИИ** в прибыли до налогообложения у Сбера ([источник](#)), что демонстрирует высокий уровень зрелости и масштаба применения технологий.

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - **Среднегодовой темп роста со сдвигом на уровень зрелости**
- Потенциальный бенчмарк:
 - 50% — **среднегодовой темп роста финансового эффекта от ИИ** в Сбере ([интервью А.А.Ведяхина](#)) (2020–2024), что свидетельствует о ускоряющемся ROI и масштабировании решений.
- Предлагаемая метрика:
 - **ROI**
- Потенциальный бенчмарк:
 - 1:8 — **окупаемость инвестиций в ИИ** в Сбере ([исследование SAP](#)), что говорит о высокой эффективности принимаемых решений об инвестициях.
- Предлагаемая метрика:
 - **Темп роста инвестиций**
- Потенциальный бенчмарк:
 - x1,5 — планы Сбера по **росту инвестиций в ИИ** до 2029 года ([интервью Г.О.Грефа](#)); 88% — средний прирост инвестиций в ИИ в Т-Банке за 2023–2025 годы (внутренний прогноз по росту инвестиций в ИИ до конца 2025г).

Эти данные подтверждают, что лидеры рынка не только активно инвестируют в ИИ, но и добиваются значительной отдачи. Компаниям важно не просто увеличивать бюджеты, но и контролировать их эффективность.

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % процессов / продуктов / бизнес-функций с ИИ
- Потенциальный бенчмарк:
 - ~80% — **доля бизнес-процессов, использующих ИИ** в Сбере ([выступление А. Ведяхина / обзор TAdviser](#)). Это показывает высокую степень интеграции технологий во все аспекты деятельности.

Ритейл и e-com

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:

- +1,2–1,9 п.п. — потенциальный прирост маржи EBIT от генеративного ИИ в ритейле; до 5% прироста продаж дают ИИ-системы поддержки решений. [\(источник\)](#)

Эффект концентрируется в мерчандайзинге, маркетинге, сервисе и цепочке поставок при перестройке процессов вокруг ИИ, а не простом добавлении инструментов.

02 Операционная эффективность

- Предлагаемая метрика:
 - % оптимизированного OPEX за счёт ИИ / ГенИИ
- Потенциальный бенчмарк:
 - ~15% — сокращение операционных затрат у ритейлеров, применяющих ИИ в масштабе; по отраслевым опросам — до 30%. [\(источник\)](#)

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - ROI
- Потенциальный бенчмарк:
 - 55% ритейлеров фиксируют ROI от ИИ выше 10%, 21% — выше 30% (отраслевые опросы 2025). [\(источник\)](#)
- Предлагаемая метрика:
 - Темп роста инвестиций компании в ИИ (год к году)
- Потенциальный бенчмарк:
 - Ориентир — опережающий рост собственного ИИ-бюджета: у ритейлеров доля ИИ в технологическом бюджете растёт быстрее самого ИТ-бюджета (в среднем по сегменту с 15% до 20% за год). [\(источник\)](#)

Метрика отслеживается по внутренней динамике компании; сравнивать корректно с собственной траекторией и с лидером сегмента.

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % собственных процессов / продуктов / бизнес-функций компании с ИИ
- Потенциальный бенчмарк:
 - Ориентир по лидеру: крупнейшие ритейлеры доводят покрытие ключевых операций ИИ до большинства — например, цель ~65% магазинов с автоматизированным фулфилментом и сквозное применение ИИ в мерчандайзинге, ценообразовании и цепочке поставок. [\(источник\)](#)

Метрика измеряет глубину проникновения ИИ внутри самой компании (доля её процессов/продуктов с ИИ), а не долю компаний отрасли, внедривших ИИ.

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:
 - Устойчивого публичного бенчмарка по доле в прибыли пока нет; в качестве ориентира используется снижение затрат (см. направление 2).

02 Операционная эффективность

- Предлагаемая метрика:
 - % оптимизированного OPEX за счёт ИИ / ГенИИ
- Потенциальный бенчмарк:
 - 5–20% — сокращение логистических затрат при встроенном в операции ИИ. [\(источник\)](#)

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - ROI
- Потенциальный бенчмарк:
 - 3,5x — медианный ROI ИИ-инвестиций в логистике за 3 года. [\(источник\)](#)

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % собственных процессов / продуктов / бизнес-функций компании с ИИ
- Потенциальный бенчмарк:
 - Ориентир — доля рейсов, маршрутов и решений диспетчеризации, оптимизируемых ИИ внутри компании. Публичных внутрикомпанийных значений мало; у лидеров ИИ встроен в ядро операций, и цель задаётся по лидеру. [\(источник\)](#)

Метрика измеряет глубину проникновения ИИ внутри самой компании, а не долю компаний отрасли с ИИ.

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:

- +3–4 п.п. за 2 года и до +8–10 п.п. за 5 лет — потенциальный прирост EBITDA-маржи и ROIC от ИИ. [\(источник\)](#)

Прирост достигается за счёт улучшения управления жизненным циклом клиента (выручка) и снижения затрат по всем доменам.

02 Операционная эффективность

- Предлагаемая метрика:
 - % оптимизированного OPEX за счёт ИИ / ГенИИ
- Потенциальный бенчмарк:
 - ~15–30% — снижение операционных затрат за счёт ИИ/ГенИИ в маркетинге, продажах и клиентском обслуживании (отраслевой ориентир). [\(источник\)](#)

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - Темп роста инвестиций компании в ИИ (год к году)
- Потенциальный бенчмарк:
 - Метрика — рост собственного ИИ-бюджета оператора; лидеры наращивают инвестиции опережающими темпами на горизонте до 2029 года. Точных публичных значений мало — ориентир задаётся по внутренней траектории. [\(источник\)](#)

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % собственных процессов / продуктов / бизнес-функций компании с ИИ
- Потенциальный бенчмарк:
 - Ориентир — доля сетевых и клиентских операций под управлением ИИ внутри компании. У лидеров автоматизирована значительная часть сетевых операций, что при масштабе даёт 30–70% меньше тикетов и 55–80% сокращения затрат центра управления сетью. [\(источник\)](#)

Метрика измеряет глубину проникновения ИИ внутри компании, а не долю операторов отрасли, фиксирующих эффект.

АПК, промышленность и энергетика

01 Финансовый эффект от ИИ

- Предлагаемая метрика:
 - % (доля) фин. эффекта от ИИ в общем фин. эффекте компании (прибыль до налогообложения)
- Потенциальный бенчмарк:
 - До 30–70% от EBIT — потенциальный прирост прибыли у компаний, полностью внедривших ИИ, на горизонте 5 лет. [\(источник\)](#)

Это верхняя оценка для лидеров полного внедрения; для большинства компаний фактический эффект на EBIT пока составляет менее 5%.

02 Операционная эффективность

- Предлагаемая метрика:
 - % оптимизированного OPEX / рост операционной эффективности за счёт ИИ
- Потенциальный бенчмарк:
 - 10–25% — рост операционной эффективности от внедрения ИИ. [\(источник\)](#)

03 Темпы роста и инвестиционная активность

- Предлагаемая метрика:
 - Темп роста инвестиций компании в ИИ (год к году)
- Потенциальный бенчмарк:
 - Метрика — рост собственного ИИ-бюджета; отдача растёт по мере перехода от пилотов к сквозной трансформации (по рынку 88% компаний применяют ИИ хотя бы в одной функции, но лишь 39% видят влияние на EBIT). Ориентир задаётся по внутренней траектории. [\(источник\)](#)

04 Уровень применения ИИ

- Предлагаемая метрика:
 - % собственных процессов / продуктов / бизнес-функций компании с ИИ
- Потенциальный бенчмарк:
 - Ориентир — доля производственных линий, активов и процессов компании, охваченных ИИ (например, предиктивным обслуживанием и оптимизацией режима). По индустрии предиктивным обслуживанием охвачено лишь 30–40% предприятий, поэтому у большинства внутренняя глубина пока низкая, а цель — расширение охвата до уровня лидеров. [\(источник\)](#)

Метрика измеряет глубину проникновения ИИ внутри самой компании, а не долю предприятий отрасли с ИИ.

Рекомендации по применению

Слепое сравнение с рыночными бенчмарками может привести к серьезным ошибкам в оценке эффективности компании. Это происходит потому, что прямое заимствование чужих метрик без учета внутреннего контекста:

- 1 Игнорирует разницу в зрелости процессов — сравнивая KPI цифрового гиганта с показателями традиционной компании, вы рискуете принять нереалистичные цели, не соответствующие текущим возможностям организации;
- 2 Не учитывает специфику бизнес-модели — банк с розничным фокусом и инвестиционный холдинг имеют принципиально разные метрики рентабельности и риска, что делает их прямое сопоставление бессмысленным;
- 3 Разная методология данных — если отраслевой бенчмарк построен на агрегированных данных по стандартизированной методике, а ваша компания использует внутренние данные с иным уровнем детализации и обработки, прямое сравнение показателей некорректно. Это подобно сравнению

точности медицинского термометра и прогноза погоды — они измеряют разное, с разной целью и погрешностью.

Использование бенчмарков должно быть комплексным и учитывать специфику компании: степень зрелости процессов, систему управления данными, бизнес-модель и отраслевые особенности. Только контекстный анализ позволяет сделать объективные выводы и избежать ошибочных решений. Представленные бенчмарки носят информационно-справочный характер и не могут рассматриваться как обязательные нормативы или целевые показатели.

- Представленные метрики можно использовать как базовые ориентиры для внутреннего анализа.
- Сравнивайте свои показатели с рыночными данными (консалтинговые отчеты, публичные данные представителей индустрии) для формирования выводов о развитии ИИ в компании.
- Отслеживайте динамику изменений: например, если текущий план развития ИИ не предусматривает достижения цели в 20% финансового эффекта от ИИ в EBIT, у компании есть потенциал для расширения ИИ портфеля и увеличения эффектов от него.
- Оценивайте не только ваше текущее положение относительно рынка, но и динамику: даже при низком текущем эффекте, высокий темп роста может свидетельствовать о правильном пути развития.

Оценка финансового эффекта от ИИ должна быть не статичной, а динамической и ориентированной на сравнение с лучшими практиками. Бенчмарки по ключевым метрикам — финансовый эффект, сокращение расходов, ROI, темп роста и глубина внедрения — позволяют выстроить стратегию, направленную на достижение лидерства в цифровой трансформации. Регулярный мониторинг и корректировка подхода на основе внешних данных с учетом внутренней специфики компании помогут не только удерживать конкурентное преимущество, но и создавать долгосрочную ценность.

Корректность расчета

Корректность оценки эффекта

ТОП-ошибок при расчете эффекта

Ошибки мышления (Cognitive Biases)

Помимо технических ошибок в расчётах, существуют систематические искажения мышления, которые влияют на оценку ещё до начала работы с цифрами. Их важно осознавать, поскольку они трудно поддаются устранению через методологию — требуется осознанная работа с командой и процессом валидации.

01 Эффект снятия сливок (Cherry-Picking)

Суть: Отбор только тех данных или сегментов, которые подтверждают позитивный эффект, при игнорировании остальных.

Проявляется двумя способами:

а) Отбор «удобного» временного окна.

Период для оценки выбирается постфактум — там, где эффект выглядит максимально. Например, ИИ-модель показала рост конверсии в течение двух недель после запуска — именно этот период берётся как репрезентативный, хотя через месяц метрика вернулась к базовому уровню.

б) Отбор «удобного» сегмента.

В A/B тест включают заранее лояльных пользователей или наиболее прибыльные сегменты. Например, модель персонализированных рекомендаций тестируется только на активных покупателях с высоким средним чеком — результат выглядит убедительно, но не масштабируется на всю базу.

Защита: фиксировать период и периметр оценки до начала эксперимента; при подведении итогов проверять результаты на полной выборке, а не только на «лучшем» срезе.

02 Моральный риск (Moral Hazard)

Суть: Стороны, заинтересованные в положительном результате оценки, осознанно или неосознанно влияют на её условия или интерпретацию — зная, что последствия завышенного эффекта лягут на других.

Проявляется в нескольких формах:

а) Команда разработки влияет на дизайн пилота.

Команда, которая создаёт модель, самостоятельно определяет контрольную группу или метрику успеха. Это конфликт интересов: результат оценивает тот, кто мотивирован показать положительный результат.

б) Бизнес-заказчик завышает базовый сценарий.

Заказчик инициативы намеренно занижает показатели «без ИИ» (базовый сценарий), чтобы дельта эффекта

выглядела больше. Например, плановый показатель конверсии без модели устанавливается на заведомо низком уровне.

в) Исполнитель оптимизирует под измеряемую метрику, а не под бизнес-результат.

Сотрудник или подрядчик, зная, что оценивается только целевая метрика (например, скорость обработки заявок), оптимизирует именно её — за счёт качества или смежных показателей, которые не измеряются.

Защита: разделять роли разработчика модели и оценщика эффекта; фиксировать базовый сценарий и метрики до запуска пилота; привлекать независимую финансовую валидацию.

03 Предвзятость подтверждения (Confirmation Bias)

Суть: Склонность искать, интерпретировать и запоминать только ту информацию, которая подтверждает уже сложившееся убеждение о ценности инициативы.

Проявляется в трёх аспектах:

а) Избирательная интерпретация результатов.

При неоднозначных итогах A/B теста команда выбирает ту трактовку, которая поддерживает гипотезу об эффективности модели, игнорируя альтернативные объяснения (сезонность, внешние события, изменение поведения контрольной группы).

б) Игнорирование отрицательных сигналов при мониторинге.

После запуска модели в прод снижение целевой метрики объясняется внешними факторами («сезонность», «рынок»), тогда как рост — заслугой ИИ.

в) Давление на аналитика.

Команда, убеждённая в ценности инициативы, неосознанно давит на аналитика с целью «найти» положительный эффект. Аналитик, в свою очередь, может варьировать методологию до тех пор, пока результат не станет «нужным».

Защита: документировать гипотезы и ожидаемый диапазон эффекта до расчёта; использовать внешнюю валидацию итогов; применять заранее зафиксированный протокол анализа.

Расчётные ошибки

01 Игнорирование расходов, связанных с инициативой

Суть: В расчёт берётся только прямой доход или экономия от ИИ, без учёта затрат на разработку, внедрение, инфраструктуру и поддержку модели.

Пример: ИИ-модель увеличивает конверсию в продажу подписки на 15%. В расчёт эффекта включается только прирост выручки. При этом не учитываются: расходы на разработку модели, стоимость облачной инфраструктуры, затраты на разметку данных и периодическое переобучение, а также расходы на коммуникацию с клиентами (рассылки, пуши), которые обеспечивают этот прирост. В итоге gross-эффект принимается за net.

Решение: при расчёте всегда использовать net-эффект — вычитать все инкрементальные затраты, связанные с инициативой (см. главу «Чистый (net) vs Валовый (gross)»).

02 Учёт виртуальной экономии как реальной

Суть: Экономия от ИИ не материализуется в реальное снижение расходов или рост прибыли — она остаётся «на бумаге».

Пример: ИИ-генератор изображений позволяет дизайнеру выполнять задачи на 30% быстрее. Компания фиксирует «экономия» 0,3 ставки дизайнера. Однако дизайнер продолжает работать полный день — высвободившееся время заполняется другими задачами или простаивает. Реального сокращения ФОТ или роста выручки не происходит.

Решение: фиксировать только те эффекты ПШЕ, которые отражены в бизнес-плане как реальные сокращения или подтверждённый найм; либо доказывать, что высвобождённое время перераспределяется на задачи с измеримым финансовым результатом.

03 Некорректный базовый сценарий (Baseline)

Суть: Эффект рассчитывается относительно неправильно выбранной «точки отсчёта» — что завышает или занижает итоговый результат.

Пример: Компания внедряет ИИ-чат-бота для обработки входящих обращений. В качестве базового сценария берётся текущая стоимость обработки одного обращения оператором. Однако без бота компания планировала нанять дополнительных операторов под рост объёма — именно этот сценарий (с учётом найма) является корректным базовым. Использование текущих, а не прогнозных расходов занижает реальный эффект.

Решение: базовый сценарий должен отражать наиболее вероятное развитие событий без ИИ — с учётом планового роста объёмов, изменений процессов и других факторов.

04 Игнорирование каннибализации и кросс-эффектов

Суть: Учитывается только прямой прирост по целевому продукту или процессу, без оценки влияния на смежные продукты, каналы и ресурсы.

Пример: ИИ-модель рекомендаций увеличивает продажи продукта А на 20%. Однако продукт Б — его функциональный аналог — одновременно теряет 15% продаж. Итоговый эффект значительно ниже, чем кажется при взгляде только на продукт А. Дополнительно: если время сотрудников или бюджет коммуникаций ограничены, рост по одному направлению всегда происходит за счёт другого.

Решение: оценивать эффект по всему «сплиту» — полке продуктов, каналам, ресурсным ограничениям — а не только по целевому объекту воздействия.

05 Игнорирование долгосрочных и отложенных эффектов

Суть: В расчёт берётся только краткосрочный эффект, без учёта изменений, которые проявляются позже — как позитивных, так и негативных.

Пример: ИИ-чат-бот снижает операционные расходы на поддержку в первый год. Однако не учитывается: деградация модели без переобучения через 12–18 месяцев, рост недовольства клиентов при снижении качества ответов, а также затраты на последующую доработку или замену решения. Расчёт даёт положительный ROI, который не реализуется на горизонте 3 лет.

Решение: строить финансовую модель на полном горизонте жизни продукта/инициативы; закладывать расходы на поддержку, переобучение и мониторинг модели.

06 Статистически незначимый результат принимается за эффект

Суть: Положительная дельта между тестовой и контрольной группой интерпретируется как доказанный эффект, хотя разница находится в пределах статистической погрешности.

Пример: A/B тест показал рост конверсии с 4,0% до 4,3% в тестовой группе. Команда фиксирует эффект +7,5%. Однако при размере выборки в 500 человек на группу p-value составляет 0,18 — результат статистически незначим. Реального эффекта нет, есть случайный шум.

Решение: до начала эксперимента рассчитывать необходимый размер выборки для достижения нужной статистической мощности; принимать результат только при достижении заранее установленного уровня значимости (как правило, $p < 0,05$).

07 Атрибуция эффекта не тому объекту

Суть: Улучшение метрики приписывается ИИ-инициативе, хотя фактически вызвано другими изменениями, происходившими одновременно.

Пример: Компания запускает ИИ-инструмент для помощи менеджерам по продажам и одновременно проводит их переобучение. Продажи вырастают на 12%. Весь эффект атрибутируется ИИ. Однако без контрольной группы невозможно разделить вклад инструмента и тренинга — реальный эффект от модели может быть нулевым или даже отрицательным.

Решение: изолировать тестируемую переменную; при невозможности проведения чистого A/B теста использовать квази-экспериментальные методы с явным описанием допущений и неопределённости (см. главу «Методы оценки эффекта»).

08 Некорректный дизайн эксперимента

Суть: Тест формально проведён, но его дизайн содержит ошибки, которые делают выводы недостоверными.

Пример: Компания тестирует ИИ-маршрутизацию для курьеров. Тестовая группа — курьеры в крупных городах с плотной застройкой, контрольная — курьеры в пригородах. Тестовая группа показывает экономию топлива на 18%. Однако разница объясняется не моделью, а принципиально разными условиями работы: плотность точек доставки, пробки, расстояния. Выборки несопоставимы изначально.

Типичные ошибки дизайна:

Несопоставимые группы — тестовая и контрольная группы отличаются по ключевым характеристикам помимо тестируемого воздействия.

Загрязнение групп (Spillover) — участники контрольной группы косвенно получают воздействие (например, сотрудники общаются и перенимают практики из тестовой группы).

Слишком короткий период — эффект не успевает стабилизироваться, фиксируется новизна или аномальный период.

Отсутствие рандомизации — группы формируются по удобству, а не случайно, что вносит систематическое смещение.

Решение: проектировать дизайн эксперимента до его начала; обеспечивать сопоставимость групп через рандомизацию или стратификацию; закладывать достаточный горизонт наблюдения

Двойная оценка фин.эффекта

Гранулярность расчета эффектов для избежания

В зависимости от поставленной цели, финансовый эффект от изменения в бизнес-процессе может рассчитываться на различных уровнях — от наибольшего макроуровня влияния на весь процесс до микроуровня влияния на конкретный компонент процесса.

Объектом финансовой оценки является внесение изменения в операционную модель бизнес-процесса. Преимущественно используются следующие уровни расчета эффектов:

- Инициатива;
- продукт/канал/отдельная ветка бизнес-процесса или технологии;
- модель.

ИИ-инициатива предполагает комплексное изменение бизнес-процесса, предполагающее внедрение/модификацию моделей, изменение критериев и настроек системы принятия решений. Как правило, ИИ-инициатива предполагает серьезное изменение бизнес-процесса.

Примеры ИИ-инициатив:

- в ритейле — внедрение динамического ценообразования: модель в режиме реального времени корректирует цену на товар в зависимости от спроса, остатков и действий конкурентов;
- в логистике — оптимизация маршрутов доставки: модель перестраивает маршруты курьеров с учётом загруженности дорог, временных окон и приоритетов заказов.

ИИ-инициатива может быть как частью одного продукта, так и предполагать внесение изменений в компоненты нескольких продуктов, каналов и прочих модальностей/технологических процессов. В то же время, ИИ-инициатива может быть частью одного продукта, но предполагать разработку группы моделей.

Выбор уровня granularity для расчета финансового эффекта зависит от специфики процесса, переводимого на модельный подход. К расчету эффекта на уровне модели крупные организации прибегают в случаях, когда инициатива по изменению процесса состоит из одной модели или же когда модель имеет высокий уровень значимости для бизнеса.

Классы потенциальных задвоений

Принято выделять пять наиболее характерных причин потенциальных задвоений финансовых эффектов:

1. Технологические и архитектурные причины: модульная многоступенчатая логика применения моделей в бизнес-процессе (модульные/каскадные модели) и сложность ее учета в оценке финансового эффекта.

Каскадные модели — это группы моделей, направленные на один и тот же объект управления (продукт, услугу, клиента). Пример: каскад моделей рекомендательной системы, последовательно фильтрующих и ранжирующих товары для одного и того же покупателя.

Модульные модели — это модели с несколькими отдельными составляющими компонентами, соединенными параллельно и/или последовательно. Пример: модель прогнозирования оттока клиентов, представляющая собой ансамбль из нескольких модулей (поведенческий модуль, модуль продуктовой активности, модуль внешних сигналов).

2. Процессные причины: реализация параллельных бизнес-инициатив несколькими командами с общим КПЭ. Пример: несколько команд независимо разрабатывают модели персонализации для разных каналов взаимодействия с клиентом (мобильное приложение, сайт, колл-центр) при общей метрике «конверсия в покупку» — каждая команда может засчитать себе финансовый эффект, влияя на один и тот же итоговый результат.

3. Методологические причины: расчёт комбинированных, косвенных и взаимозависимых метрик. Комбинированная метрика — это метрика, зависящая от нескольких факторов. Пример: LTV клиента, где часть эффекта зависит от модели удержания, часть — от модели кросс-продаж; изменение одной модели автоматически влияет на значение общей метрики, что может создать иллюзию двойного эффекта. Таким образом, выбор неверной метрики может отражать финансовый эффект не только от модели, но и от всего бизнес-процесса.

4. Операционные причины: повторное применение одной модели/сервиса в разных процессах или каналах. Пример: модель оценки платёжеспособности клиента используется одновременно в процессе первичной продажи и в процессе реактивации ушедших клиентов — финансовый эффект от одной и той же модели может быть учтён дважды в разных инициативах.

5. Расчет эффекта от новой версии модели в текущем процессе без учета инкрементальности эффекта. Эффектом новой версии модели признается не очередной инкремент, который она вносит в процесс, а общий эффект по сравнению с отсутствием модели в целом, что приводит к существенному завышению результатов.

Способы ведения системного учета влияния ИИ-моделей на бизнес-процессы организации

Системный учет влияния ИИ-моделей на бизнес-процессы — инструмент, позволяющий повысить качество управления модельным риском, в том числе в части контроля ситуаций возникновения удвоения расчета финансового эффекта.

Элементами системного учета являются:

- формализованный и утвержденный на уровне всей организации подход и методика оценки фин. эффектов ИИ с закрепленным распределением ролей участвующих в оценке подразделений;
- реестр бизнес-процессов компании с маркировкой подпроцессов, переведенных на модельный подход, едиными уникальными идентификаторами ИИ-моделей, работающих в процессах;
- реестр финансовых, продуктовых и математических метрик на принятом в компании уровне гранулярности расчета эффектов с маркировкой потенциального модельного риска; предполагается учет как потенциальных метрик, рассчитанных при возникновении идеи о реализации инициативы, так и фактических, полученных на стадиях от прототипирования до внедрения и регулярной валидации;
- реестр бизнес-документации по ИИ-инициативам;
- реестр технической документации по ИИ-моделям;
- иные артефакты по моделям ИИ, являющиеся обязательными с точки зрения внутренних регламентных процедур организации.

Методика оценки двойного эффекта

Инструменты разделения вклада ИИ-инициатив и избегания двойного учета

1. Для количественного разделения вклада инициатив на уровне моделей применяются методы множественного А/Б тестирования, как последовательного, так и параллельного. Процесс А/Б тестирования подробно описан в главе «Какими методами можно посчитать эффект от ИИ».
 - В случаях, когда несколько моделей влияют на одного клиента/сотрудника, используется техника разведения пилотов во времени для того, чтобы изолировать вклад каждой модели.
 - В случаях, когда одна модель используется в нескольких процессах, рассчитывается эффект от применения модели в каждом процессе на уровне гранулярности, принятом в компании.
 - Если же требуется оценить эффект от «типа модели», то эффект рассчитывается как сумма эффектов на уровне модели.
2. Для операционного разделения вклада инициатив используются инструменты ведения системного учета, в том числе присвоение ИИ-моделям уникальных финансовых идентификаторов с занесением в реестры бизнес-процессов и метрик.
3. В сложных случаях, в том числе при предпосылках каннибализации или отсутствии экономической целесообразности в проведении множественного тестирования, используются:
 - процедуры укрупнения расчетов на уровень выше, к примеру, с уровня модели до каскада моделей, с уровня инициативы до уровня продукта и т.д.;
 - процедуры приоритизации, к примеру, эффект может засчитываться основной самой доходной модели в инициативе/продукте/каскаде.

Примеры:

- для сложных процессов, где каждый этап покрыт несколькими взаимосвязанными моделями, рассчитываются эффекты в двух разрезах: по каждой из моделей в отдельности и общий эффект от всего комплекса ИИ-внедрений на процесс. Это позволяет учесть пересечения эффектов и корректно оценить суммарный прирост;
 - для составных моделей, включающих несколько компонентов с различной архитектурой, рассчитывается эффект от каждого компонента, а также суммарно от всей модели целиком;
 - при обновлении текущей версии модели в эффект учитывается только инкремент по сравнению с предыдущей версией, что позволяет корректно оценить влияние дополнительных трудозатрат в привязке к дополнительной ценности.
4. Для контроля учета используется независимая экспертная оценка проводимых расчетов с целью выявления потенциальных процессов и зон, подверженных двойному учету финансовых эффектов от внедрения моделей.

В каких случаях не учитывается эффект от отдельно взятых моделей

Существует четыре типа ситуаций, в которых расчет финансового эффекта не производится:

- если эффект незначительный;
- если недостаточно данных для корректной оценки вклада инициативы;
- если модель не дает самостоятельного результата (в таком случае используются методы укрупнения);
- если самостоятельный результат отдельно взятой модели не представляет интерес для бизнеса.

Финансовая индустрия

В финансовых организациях наиболее подвержены риску двойного учёта эффекта процессы с множественным применением моделей, смешения с бизнес-правилами, процессы с общими КПЭ, а также процессы, где симуляция работы процессов без ИИ-модели невозможна. Наиболее типичные области и кейсы:

ТИП ПРОЦЕССА	ПРИМЕР
Маркетинг (склонность, удержание, акции)	Модель склонности к отклику и модель удержания одновременно влияют на одного и того же клиента
Кредитование и скоринг (PD модели)	Модель PD и скоринговые модели влияют на одни и те же кредитные решения
Ценообразование и дисконтная политика	Модели ценообразования по кредитам, депозитам, кэшбэку одновременно влияют на LTV клиента
Программы лояльности и клиентский опыт	Модель прогнозирования оттока и модель прогноза кросс-продаж влияют на одну и ту же дополнительную выручку

Примеры:

- модели продуктовых склонностей используются финансовыми организациями во всех бизнес-линиях. При этом оценка финансового эффекта на уровне одной модели сопряжена с большими затратами на изоляцию воздействия на клиента исключительно одной модели при сохранении остальных критериев неизменными. В данном случае оценка происходит на уровне выше - на уровне алгоритма оптимизации (использующего все остальные модели склонности) и предложения лучшего продуктового оффера клиенту;
- модели оценки вероятности дефолта имеют многомодульную разветвлённую структуру и обращаются при оценке кредитоспособности клиента к множеству источников информации. Оценка финансового эффекта каждого отдельного модуля модели не производится, однако производится оценка и пилотирование эффекта на уровне кредитного процесса с точки зрения влияния на выдачи, уровень риска и пр. бизнес-метрики.

Ритейл и e-com

В ритейле и e-com риск двойного учёта возникает там, где на сквозные показатели продукта (товарооборот, конверсия, средний чек) одновременно работает множество моделей, а также в процессах с общим ресурсом (курьеры) и в маркетплейс-среде, где воздействие на одну часть аудитории меняет поведение другой. Наиболее типичные области и кейсы:

Примеры:

- рекомендации, ранжирование выдачи, персонализация и ценообразование одновременно влияют на конверсию и средний чек одного и того же пользователя. Сумма эффектов отдельных А/Б по фичам систематически не сходится с общим приростом товарооборота, поэтому совокупный эффект признаётся

на уровне всего продукта через постоянную (глобальную) контрольную группу, а не суммированием отдельных тестов;

- в доставке и q-commerce курьеры — общий ресурс: ускорение или удешевление доставки для одной группы заказов меняет доступность курьеров для другой, из-за чего эффекты отдельных моделей перетекают между группами. Изоляция вклада модели требует switchback-дизайна, а итоговый эффект считается на уровне процесса;
- промежуточные пользовательские метрики (добавления в корзину, просмотры) часто улучшаются несколькими моделями сразу — если переводить в деньги каждую по отдельности, возникает двойной учёт. Денежный эффект признаётся один раз на верхнем сквозном показателе.

Транспорт и логистика

В транспорте и логистике риск двойного учёта концентрируется в процессах с общим ресурсом и сетью (распределение мест, бортов, экипажей, курьеров) и в сквозных процессах доставки, где на один КПЭ работает каскад моделей. Наиболее типичные области и кейсы:

Примеры:

- на единый КПЭ доставки (доля доставок вовремя, стоимость доставки) одновременно влияют прогноз спроса, расчёт времени прибытия (ETA), оптимизация маршрутов и динамическая диспетчеризация. Их эффекты перекрываются, поэтому суммирование оценок отдельных моделей завышает результат; эффект замеряется на уровне сквозного процесса доставки, а вклад отдельной модели — только в изолированном эксперименте;
- ценообразование и распределение ресурсов — процессы с сильной интерференцией: воздействие на одну группу (слот, маршрут, регион) меняет показатели другой через общий рынок и общий парк. Оценка ведётся на уровне процесса, а изоляция модели требует switchback или синтетического контроля;
- на критических процессах, где работу без модели невозможно воспроизвести, эффект признаётся на уровне процесса, а не отдельной модели.

ИКТ

В телекоме риск двойного учёта концентрируется в процессах удержания и монетизации абонентов, где работает каскад пересекающихся моделей на общие КПЭ (отток, выручка на абонента), и в мультисегментной структуре, где одна инициатива затрагивает несколько сегментов и продуктов. Наиболее типичные области и кейсы:

Примеры:

- удержание абонентов реализовано как непрерывный процесс с каскадом моделей (сегментация и таргетинг, прогноз оттока, выбор следующего действия, динамические предложения), работающих на общие показатели — отток и выручку на абонента. Из-за высокого пересечения аудитории эффект отдельной модели почти не выделяется; оценка и пилотирование ведутся на уровне всего процесса удержания;
- выбор следующего действия (next best action) и рекомендации тарифов часто работают поверх тех же абонентов, что и модель оттока, — их эффекты нельзя складывать; признаётся совокупный эффект на уровне алгоритма принятия решения;

- при пакетной продаже (связь, интернет, ТВ) и в разных сегментах (B2C / B2B / HR) одна модель влияет на несколько продуктов. Денежный эффект признаётся по всей продуктовой корзине один раз, без повторного учёта в каждом продукте.

АПК, промышленность и энергетика

В этих отраслях риск двойного учёта концентрируется на производственных процессах, где несколько моделей одновременно влияют на один интегральный показатель (например, общую эффективность оборудования — ОЕЕ, выход годной продукции, удельные затраты), а остановить производство для чистого сравнения «с моделью / без модели» невозможно. Наиболее типичные области и кейсы:

Примеры:

- на одной производственной линии предиктивное обслуживание, оптимизация технологического режима и автоматический контроль качества бьют в один показатель эффективности: предиктивное обслуживание снижает аварийные простои, оптимизация режима — скоростные потери, контроль качества — брак. Так как эффекты сходятся на общем КПЭ, их раздельное суммирование завышает результат; эффект признаётся на уровне линии или технологического контура;
- в нефтегазе и энергетике модели оптимизации режима добычи и выработки, прогноза отказов и планирования ремонтов работают на общие показатели — объём добычи и удельные затраты. Вклад отдельной модели выделяется только при наличии нескольких сопоставимых агрегированных единиц (через синтетический контроль); иначе эффект признаётся на уровне актива;
- натуральные улучшения (рост коэффициента технической готовности, снижение брака) переводятся в деньги один раз и на верхнем уровне процесса. Повторный учёт одного и того же физического улучшения в эффекте нескольких моделей недопустим.

Процессы, наименее подверженные риску двойного учета.

Существует класс процессов, в котором риски возникновения двойного учета при расчёте финансового эффекта находятся на низком уровне. Такие процессы характеризуются присутствием одиночно функционирующих моделей с низким уровнем взаимодействия с другими моделями или ограниченным влиянием на общие КПЭ. Примерами таких процессов и моделей могут служить:

- скоринг резюме кандидатов;
- распознавание документов;
- биометрическая идентификация.

Иными словами — процессы с минимальным использованием ИИ, в которых модели работают независимо.

В вышеуказанных процессах, как правило, используется ограниченное число моделей, и логика их изоляции в процессе является достаточно очевидной.

Пример: в скоринге резюме возможно выделить целевую группу, использующую результаты модели для поиска, и контрольную группу, работающую по стандартному процессу. В этом случае оценка финансового эффекта может быть получена относительно легко и без риска получения двойного учета

Методы

Какими методами можно посчитать эффект от ИИ?

УРОВЕНЬ	МЕТОДЫ	КОГДА ПРИМЕНЯТЬ
АФИИНА I Краткосрочный финансовый	Базовый А/Б тест	Есть возможность случайного разделения на группы; достаточный объем трафика; горизонт 30–60 дней
	Альтернативное моделирование (AS IS / TO BE)	А/Б тест невозможен; ИИ уже запущен на всей аудитории; эффект через изменение процессов
АФИИНА II Среднесрочный стратегический	Расширенный А/Б тест (стратификация, CUPED, Switchback)	Нужна повышенная чувствительность; малый ожидаемый эффект; нарушение SUTVA в стандартном тесте
	Вечный А/Б тест / Holdout	Непрерывный мониторинг эффекта после промышленного запуска; отслеживание деградации модели
	Квази-эксперименты (PSM, Synthetic Control)	Случайное разделение невозможно; B2B-сегмент; ИИ уже запущен; мало контрольных объектов
	Альтернативное моделирование (AS IS / TO BE)	Портфельная оценка с косвенными эффектами на смежные процессы
АФИИНА III Долгосрочный системный	Альтернативное моделирование на исторических и рыночных данных	Системный долгосрочный эффект; оценка на уровне компании или отрасли; горизонт 3–5+ лет

При замере эффекта важно оценивать не просто корреляции между метриками ИИ-инициативы и бизнес-метриками компании, а именно **доказать прямое воздействие** ИИ-инициативы на финансовые показатели.

Выбор метода оценки зависит от задачи, горизонта и зрелости ИИ-практики в компании. На уровне 1 достаточно базовых инструментов; на уровне 2 появляются более сложные методы, позволяющие работать с косвенными эффектами и портфелем; на уровне 3 оценка строится на исторических и рыночных данных без контрольной группы.

АФИИНА I

I.1. Базовый А/Б-тест

А/Б-тест — наиболее приоритетный и точный способ оценки экономического эффекта. А/Б тестирование представляет собой метод оценки влияния целевого воздействия на процесс, путем сравнения двух

выбранных подгрупп (А – Целевая Группа – с воздействием ИИ; Б – Контрольная Группа – без воздействия ИИ).

Проведение А/Б-теста состоит из **нескольких важных этапов**:

- 1 Определение объекта воздействия, объекта рандомизации, ключевого драйвера.
- 2 Проверка качества сплита, формирование итоговой гипотезы, расчёт минимальной достаточной выборки и выбор корректного статистического критерия.
- 3 Запуск теста и выставление критериев мониторинга эффекта.
- 4 Подведение итогов теста по выставленным критериям.

Объект воздействия

Объект воздействия — это тот элемент процесса, характеристики и/или количество которого может изменяться в результате внедрения ИИ модели.

Примеры: пользователь/клиент, обращение/диалог, заказ, сотрудник/команда, счет/портфель, магазин/регион, сквозной процесс, и т.д.

Принцип выбора:

- Необходимо выбирать **наиболее верхний уровень процесса**, на котором проявляется **значимая часть изменения объекта**
- Если выбранный локальный уровень процесса не захватывает все ключевые эффекты воздействия, необходимо расширить рамки анализа до **процесса более высокого порядка** (с уровня обращения до уровня пользователя; с уровня конкретной операции до уровня бизнес-процесса, с конкретного продукта до уровня корзины продуктов и т.д.).
- При этом стоит учитывать, что чем шире определение объекта, тем сложнее будет определить воздействие. Необходимо руководствоваться принципом выбора наименее **достаточного** уровня

Примеры.

- **Чат-бот поддержки.**
Узко: обращение (время оператора, сэкономленное на одно обращение).
Шире и корректнее: **пользователь** — также учитываем влияние на частотность обращения, метрики удержания пользователя в канале.
- **Компания продаж.**
Узко: продукт (конверсия по продукту);
Шире и конкретнее: **клиент** – также учитываем конверсию по прочим продуктам корзины, частоту покупок, эластичность и т.д.

Объект рандомизации

Объект рандомизации – минимальный элемент процесса с целевым ИИ-изменением, который **случайно** назначается в контрольную/целевую группы. Как правило, объект рандомизации равен объекту воздействия, однако в некоторых случаях после определения объекта воздействия требуется провести дополнительное исследование процесса, чтобы выбрать отдельный объект рандомизации.

Принцип выбора:

- Начинаем с **объекта воздействия**, выбранного в предыдущем пункте

- Проверяем на соблюдение **SUTVA**, предполагая, что объект воздействия равен объекту рандомизации
- В случае несоблюдения необходимо разделять объекты и поднимать уровень объекта рандомизации: **кластер**. Пример: **сотрудник** -> **подразделение**, **магазин** -> **город** (все магазины в городе), **клиент** -> **гео-зона** (все клиенты в гео-зоне). Проверяем **фиксированность назначения**: объект рандомизации и соответственно объект воздействия должен на протяжении всего А/Б теста находиться только в **одной** группе: **КГ** или **ЦГ** – не должно быть переходов из одной группы в другую.
- Критерий включения в тест – четко фиксируем объекты, которые могут попасть в эксперимент; исключаем конфликты с параллельными тестами (если возможно), а также конфликты в рамках нормативных документов, принятых в компании (например, политики коммуникации, кредитная политика или же принципы запретов на дискриминацию сотрудников).

Примечание:

SUTVA (*Stable Unit Treatment Value Assumption*) — допущение, что:

- 1 **Нет интерференции объектов**: исследуемое воздействие от ИИ на один конкретный объект **не влияет** на исход другой.
- 2 **Нет скрытых версий исследуемого воздействия**: механизм воздействия от модели ИИ одинаков для всех объектов

Примеры необходимости отделения объектов рандомизации от объектов воздействия.

- **Увеличение эффективности сотрудников**: при внедрении новой системы мотивации, объект воздействия – сотрудник, однако, если разбиение на КГ и ЦГ происходит на уровне сотрудников, может случиться, что внутри одного подразделения, или, еще хуже – внутри одной команды, будут сотрудники с воздействием и без него; такая ситуация может сильно повлиять на потенциальный результат исследования, поскольку существует эффект «сарафанного радио», и такие сотрудники из КГ будут недовольны «несправедливостью», что ухудшит показатели А/Б теста и завысит дельту по метрике ключевого драйвера между ЦГ и КГ)
- **Ценообразование**: увеличили цены в некоторых магазинах в ЦГ, может вырасти спрос в соседних магазинах в КГ (при этом он не должен был измениться от целевого воздействия) из-за общего рынка и эластичности спроса
- **Алгоритм назначения сотрудников**: на часть пользователей раскатали новый алгоритм назначения курьера или сотрудника поддержки, однако совокупность сотрудников и механизм назначения общие, из-за чего изменилось время ожидания и в контрольной группе.
- **Неделимость** некоторых процессов компании: при выборе процесса, как объекта воздействия, для некоторых критических процессов невозможно подобрать случайным образом пару, чтобы разделить их на две сравнимые группы.

Ключевой драйвер

Ключевой драйвер (Ключевой показатель эффективности/Ключевая метрика) — основная метрика эффекта на выбранном объекте рандомизации, по которой принимается решение о финансовой-экономической эффективности инициативы. Должна отражать **полный эффект**, а не удобный фрагмент (см. раздел принципы финансовой оценки).

Принципы выбора:

- 1 **Полнота экономики**. необходимо учитывать все инкрементальные **доходы и затраты**, на которые потенциально может повлиять модель.

Пример: модель оптимизирует продажи продукта А для всех клиентов, однако в то же время можем

повлиять на продажи продукта Б, С, Д и т.д. – тогда необходимо так же учитывать эффект во всей линейке продуктов (каннибализация и кросс-продажи). Ключевыми драйверами в такой ситуации должны быть выбраны продуктовые метрики всех связанных продуктов полки продаж.

- 2 **Единица расчета.** На формулу расчета ключевой метрики не должно оказывать влияние исследуемое изменение от внедрения ИИ. Метрику необходимо рассчитывать на фиксированной популяции объектов, которые участвуют в тесте, не смещая и не искажая расчет. Воздействие в целевой группе часто меняет число событий (диалогов, показов, доставок, и т.д.), поэтому метрики, взятые в отношении «на событие» становятся зависимыми от воздействия. Правильный подход — считать **на единицу** объекта рандомизации.

Пример: можем считать конверсию по показам, однако ИИ модель в том числе может влиять на количество просмотров, что может исказить расчет, в таком случае лучше всего считать клики на юзера за период.

- 3 **Окно агрегирования** — фиксируем **измерительное окно ключевой метрики** (7/30/90 дней и т.п.), где материализуется значимая часть эффекта. Необходимо выбирать **минимально достаточное** окно по задержки эффекта во времени и мощности теста.
 - 4 **Направление и пороги.** Четко задаем направление улучшение метрики: «выше лучше/ниже лучше»; заранее фиксируем **MDE** (минимально полезный эффект) и бизнес-порог принятия решения.
 - 5 **Простые формулы вместо «индексов».** По возможности необходимо использовать разницу средних базовых величин (деньги/время/счетчики), избегать доли от долей и сложных агрегаций.
 - 6 **Борьба с манипуляциями и защитные метрики.** Ключевая метрика не должна стимулировать вредные практики, так же необходимо добавить 2-5 защитные метрики.
- Борьба с манипуляциями подразумевает совокупность мер, направленных на борьбу с искажениями и манипуляции данными. Включает в себя фиксацию групп и метрик на протяжении всего теста, фиксацию протокола проведения теста и невозможность менять «правила игры» по ходу А/Б теста.
 - Защитные метрики – это метрики, измеряющие потенциальные потери, которые может понести компания в ходе тестирования воздействия от ИИ. Необходимо четко прописать алгоритм принятия решения о досрочной остановке А/Б теста и пороги значимости потерь, для случая, если тест будет наносит вред процессу/финансовому состоянию/клиентам компании.

Проверка качества сплита

После корректного **выбора объекта воздействия, единицы рандомизации и ключевого драйвера** необходимо убедиться, что сплит на группы выполнен без систематических перекосов. В противном случае выводы эксперимента **не будут достоверными**, что сделает невозможным их тиражирование и практическое применение для получения финансового эффекта.

Что необходимо проверять:

- **Несоответствие долей выборки (SRM)** – проверка соответствие фактических долей контрольной и целевой групп плановому теоретическому распределению. Проверку необходимо проводить как по всем тестовым выборкам, так и по ключевым срезам (платформа, версия, регион).
Пример: Если тест был спланирован с предполагаемым распределением 50/50, а после запуска размеры групп статистически значимо отличаются (например, 60/40), это явный признак некорректного проведения эксперимента.

Баланс ковариат – проверка на сопоставимость контрольной и тестовой групп по заранее заданному набору признаков (признаков однородности объекта рандомизации): возраст, регион, платформа, пре-метрика и т.д.

Важно: в качестве признаков нельзя брать какие-либо пост-тритмент метрики, который могут теоретически измениться после применения воздействия. В качестве признаков необходимо брать свойства/метрики, которые **известны, что до запуска и не меняется из-за воздействия**: прошлое поведение клиента (покупки/клики/обращения за n дней до теста), стабильные свойства объекта рандомизации (регион, тип клиента, устройство/платформа, тариф) и контекст попадания в тест (канал, день недели/время суток и т.д.). Для определения всех релевантных признаков однородности, требуется изучить бизнес-процесс, в который внедряется ИИ, а также объект рандомизации, который был выбран для проведения А/Б теста.

А/А тест, который также должен выступать инструментом проверки качества сплита и сравнимости групп тестирования. А/А тест подразумевает собой полный аналог А/Б теста, но без целевого воздействия в одной из групп. Как правило, используется А/А тест для проверки однородности собранных групп в том числе и по ключевой метрике (ключевому драйверу) к моменту старта А/Б теста, перед запуском А/Б теста.

- **Фиксированность назначения** – не должно быть объектов рандомизации, которые находятся одновременно в контрольной и целевой группах во время теста, или могут переходить из одной группы в другую по ходу теста – **нет переназначений**

Когда проверяем:

- **Перед запуском** (предварительный запуск на 1–5% трафика) — убедиться, что разделение по группам выполнено корректно.
- **Каждый день в ходе теста** — автоматизированный мониторинг показателей по всем заданным срезам данных с системой оповещений.
- **Финальная проверка** при подведении итогов теста

Финальная гипотеза, стат. критерий и размер выборки

Далее необходимо сформулировать итоговую проверяемую гипотезу и выбрать стат. критерий.

Гипотеза в рамках А/Б теста формируется в терминах дельты (инкремента) ключевой метрики замера эффекта относительно выбранного бейзлайна. В качестве бейзлайн должна быть выбрана лучшая ближайшая альтернативу процессу TO BE.

Гипотеза по ключевому драйверу/ключевой метрике формулируется в формате:

$$H_0: \Delta = 0, H_1: \Delta \neq 0$$

По умолчанию используем двустороннюю проверку гипотезы (дельта эффекта может быть как положительной, так и отрицательной). Односторонняя проверка может быть использована, если заранее доказано, что обратный эффект невозможен.

Рекомендации по стат. критериям:

- 1 Для непрерывных метрик по умолчанию используется Welch t-test (необходима независимость наблюдений)
- 2 Для доли/бинарных распределений – z/χ^2 критерий.
- 3 Для ненормальных распределений возможно использовать ранговые критерии и методы, например U-test (однако стоит учитывать специфику проверяемой гипотезы данным критерием)
- 4 В случае метрик-отношения: дополнительно применяется Delta-метод или Линеаризация
- 5 В случае нестандартных агрегаций (медиана, квантиль) или неприменимости критериев, описанных выше, возможно применение **Bootstrap**

Далее формулируем статистическую (далее – стат.) значимость и мощность эксперимента. Базовая рекомендация – **уровень стат. значимости 5%, мощность 80%**. Данные параметры могут быть изменены

основываясь на толерантности компании к ошибкам I и II рода, однако не рекомендуется выбирать менее строгие значения параметров.

Если гипотез в рамках одного теста несколько, то необходимо делать поправки на множественное сравнение. В зависимости от конкретного случая, это может быть сделано, например, поправкой Бонферрони/методом Бонферрони-Холма/и т.д.

Размер выборки и длительность эксперимента

Основываясь на параметрах ключевого драйвера, которые были определены ранее необходимо рассчитать размер минимальной достаточной выборки и длительности теста.

Размер выборки будет в первую очередь зависеть от выбранного уровня стат. значимости, мощности, MDE и дисперсии ключевой метрики.

Формула расчета размера выборки в базовом кейсе:

$$N \approx \frac{(z_{1-\alpha/2} + z_{\text{power}})^2}{\text{MDE}^2} \times \left(\frac{\sigma_1^2}{s} + \frac{\sigma_2^2}{1-s} \right)$$

При заданных параметрах $\alpha=5\%$, $\text{power}=80\% \Rightarrow z_{1-\alpha/2} \approx 1.96$, $z_{\text{power}} \approx 0.84$

s – доля целевой группы в множестве (КГ+ЦГ), σ_1^2 – дисперсия ЦГ, σ_2^2 – дисперсия КГ.

По вышеописанной формуле определяется количество наблюдений, требуемое, чтобы доказать достижение значения MDE по ключевой метрике, и что найденная дельта ключевой метрики действительно вызвано ИИ воздействием, а не является «шумом» наблюдения в рамках дисперсии.

Таким образом, $N_{\text{ЦГ}} = sN$ и $N_{\text{КГ}} = N - N_{\text{ЦГ}}$.

Зная число наблюдений, которое можно собрать в один месяц, можно рассчитать требуемое число месяцев для подтверждения эффекта на А/Б тесте.

Важно также отметить, что точная формула расчета N может отличаться в зависимости от выбранного дизайна и критерия теста.

Далее необходимо определить длительность теста на основе размера трафика, требуемого минимального размера выборки, а также срока вызревания ключевой метрики, если срок вызревания отличается от минимального требуемого срока, рассчитанного по формуле. При определении длительности также стоит учитывать эффект новизны и сезонность. Например, необходимо запускать тест минимум на одну полную неделю, если есть недельные паттерны в данных, даже если необходимый размер выборки набирается за меньшее количество дней. Рекомендацией здесь может служить следующее: для пилотирования ИИ-инициативы и проведения А/Б теста необходимо брать срок, полученный через расчет MDE + срок вызревания эффекта (окно агрегирования).

I.II. Альтернативное моделирование

Приблизительная, экспертная оценка экономического эффекта на основе бизнес-моделирования. Строится модель «как должно работать» и рассчитывается порядок величины эффекта при заданных допущениях. Не

является квази-экспериментом и доказательство причинного влияния.

Когда применяется.

- Стоит применять только в крайнем случае: рандомизация и квази-эксперименты невозможны (правовые/этические ограничения, сильная интерференция, редкие события, очень малые выборки)
- МЛ-инициатива потенциально не приносит значимую часть от общего эффекта вклада ИИ в компании

Пример построения

- **Базовый сценарий:** прогноз без ИИ (сезонность, тренды).
- **Механика эффекта:** явная цепочка «применение МЛ → изменение метрики → деньги».
- **Сценарии и неопределенность:** Base/Low/High + при необходимости Монте-Карло симуляция по ключевым неопределенностям.
- **Юнит-экономика и ограничения:** Выручка/капитальные расходы/операционные расходы, комиссии, операционные бутылочные горлышки.
- **Валидация здравым смыслом:** бенчмарки по прошлым запускам/рынку, консистентность с PnL.

I.III. Отраслевая специфика

На уровне 1 выбор метода определяется тем, можно ли в принципе провести честный А/Б-тест: в отраслях с частотными объектами это базовый А/Б, в отраслях с неделимым производством и длинным циклом — альтернативное моделирование.

Финансовая индустрия

Метод: Базовый А/Б-тест.

Почему: Объекты воздействия — клиенты, заявки, транзакции — частотны и однородны, набираются в нужном объеме, а случайное разделение на целевую и контрольную группы не ломает процесс. Поэтому уже на уровне 1 доступен самый точный метод. Базовый А/Б применяют для оценки отдельных моделей (продуктовые рекомендации, антифрод, модели склонностей) на горизонте 30–60 дней.

Ритейл и e-com

Метод: Базовый А/Б-тест.

Почему: Заказы и пользователи массовы, платформа экспериментов позволяет регулярно проверять гипотезы, поэтому эффект отдельной фишки измеряется базовым А/Б на уровне пользователя или заказа. Уже здесь важно учитывать, что часть инициатив не двигает товарооборот напрямую — в таких случаях ключевым драйвером выбирают промежуточную пользовательскую метрику (конверсия, время обработки, частота заказов).

Транспорт и логистика

Метод: Базовый А/Б в цифровых продуктах; альтернативное моделирование — в критических процессах.

Почему: Там, где объект воздействия — пользователь приложения или цифрового сервиса (рекомендации, персонализация), применим базовый А/Б. Но процессы, связанные с безопасностью и неделимостью (расписание, загрузка бортов, работа экипажей), рандомизировать нельзя — для них уже на уровне 1 используется альтернативное моделирование AS IS / TO BE.

АПК, энергетика и промышленность

Метод: Альтернативное моделирование AS IS / TO BE.

Почему: Производство непрерывно и неделимо — разделить его на сравнимые целевую и контрольную группы невозможно. Поэтому базовым методом уже на уровне 1 становится моделирование «как было / как стало»: строится сценарий без ИИ и считается порядок эффекта. Натуральные показатели при этом переводятся в деньги. Базовый А/Б остаётся применим только во вспомогательных цифровых процессах (документооборот, HR, клиентские сервисы).

ИКТ

Метод: Базовый А/Б + предварительная оценка на этапе требований.

Почему: Абонентская база массовая, поэтому базовый А/Б применим напрямую. На самом старте, когда данных ещё нет, эффект оценивают предварительно (гипотеза и альтернативное моделирование на этапе требований), а затем подтверждают базовым А/Б. Применимость отдельных инструментов при этом сужена требованиями к используемому ПО и законодательству.

АФИИНА II

III.1. Усиленный А/Б тест

Усиленный А/Б тест строится на той же базовой логике, что описана в уровне 1, но дополняется техниками, повышающими чувствительность и надёжность оценки.

Возможна ситуация, когда на этапе дизайна требуемый срок проведения оказывается неоправданно большим. В таком случае процесс оценки нарушает принцип материальности: затраты на проведения пилота оказываются сопоставимы, или, в крайних случаях, даже больше потенциального дохода от самой ИИ модели.

В данном случае стоит обратиться к **методам ускорения А/Б-тестов:**

1. Стратификация

Стратификация — это метод снижения дисперсии. Его суть в том, что популяцию **заранее разбивают** на однородные группы (**страты**) по ключевым доэкспериментальным признакам, а затем проводят рандомизацию внутри каждой группы. При анализе результаты **взвешенно объединяются** с учетом размера каждой страты. Такой принцип **«сравнения похожего с похожим»** позволяет исключить из оценки межгрупповую дисперсию, оставив лишь дисперсию внутри групп. **В результате** требуемый объем выборки или длительность теста уменьшается.

Принцип работы:

1. Необходимо выбрать 1-3 сильных признака объекта рандомизации независимых от эффекта воздействия (например, регион или платформа)
2. Задать веса равным долям трафика на пред-тестовых данных
3. Рандомизировать объекты внутри каждой страты по плановой доле КГ/ЦГ
4. При анализе использовать взвешенную метрику

$$\hat{Y}_{strat} = \sum_{k=1}^K \pi_k \bar{Y}_k, \pi_k = \frac{N_k}{N},$$

где $\bar{Y}_k$ – среднее в страте K

Когда применять.

Внутри популяции объектов воздействия есть четкие группы с отличающейся ключевой метрикой (ОС, страны/регионы, каналы). Чем больше различия между стратами и стабильнее распределение трафика, тем выше польза.

Важно:

- 1 Не нужно чрезмерно дробить трафик, необходимо держать разумный размер страты, редкие категории объединять в «other»
- 2 Страты необходимо фиксировать до запуска и проверять SRM по ним

02 CUPED (ковариатная коррекция доэкспериментальной метрикой)

CUPED - метод снижения дисперсии за счет вычитания из итоговой метрики части дисперсии, объясняемой пред-тестовой метрикой (значением метрики до старта теста или другой инвариантной к тестовому воздействию величиной).

Принцип работы:

Берем итоговую метрику Y и пре-метрику X на том же объекте рандомизации (пользователь/клиент/зона), считаем преобразованную метрику

$$Y^* = Y - \vartheta(x - \bar{x}), \theta = \frac{\text{Cov}(Y, X)}{\text{Var}(X)}$$

Далее сравниваем Y* между тестом и контролем обычным способом (Welch t-test).

Дисперсия новой преобразованной метрики $\text{Var}(Y^*) = \text{Var}(Y)(1 - \rho^2)$, где $\rho = \text{corr}(Y, X)$.

Соответственно, чем больше корреляция ковариаты и изначальной метрики, тем эффективнее будет работать метод.

Когда применять

Есть явная стабильность поведения объектов воздействия во времени: прошлые покупки, частота обращений, средний чек и т.п.

Важно: X должна быть строго инвариантна к эффекту воздействия теста. В идеале это должна быть метрика, рассчитанная до попадания в тест.

Вечный А/Б тест с глобальной контрольной группой

Глобальная контрольная группа (ГКГ) — это постоянная группа, формируемая из подходящей популяции объектов, которая не получает новых МЛ-моделей (либо получает только стабильную базовую версию). ГКГ помогает отслеживать долгосрочный дрейф данных и поведения объектов, а также выявлять потенциальную деградацию модели. Ее использование позволяет непрерывно контролировать эффект в рабочей среде и более корректно оценивать вклад множества параллельных экспериментов. ГКГ по своей сути является непрерывным А/Б-тестом. Метод стоит рассматривать как дополнение к разовым А/Б-тестам, а не как замену. ГКГ позволяет более точно и честно оценить интегральный экономический эффект, но требует значимых усилий в реализации и поддержке.

При решении о создании ГКГ стоит руководствоваться принципом материальности, в большинстве случаев затраты на оценку могут превышать потенциальную пользу. Рекомендуется использовать ГКГ для **крупных МЛ-продуктов**, которые приносят значимую долю от общего экономического эффекта от ИИ в компании и/или для МЛ-продуктов с выделенной командой, которая на постоянной основе улучшает модель.

Дизайн:

- 1 **Состав.** В ГКГ входят те же объекты, удовлетворяющие условию включения, что и в разовый А/Б-тест.
- 2 **Размер.** Обычно 1–5% трафика, стоит ориентироваться на тот же MDE, что в разовом А/Б. Доля фиксируется заранее.
- 3 **Фиксированность назначения.** Один и тот же объект всегда остается в ГКГ (если не было намеренного перемешивания), не должно быть переназначений.
- 4 **Логика сплита.** Та же, что и в разовом А/Б-тесте. Важно, чтобы новые пользователи (объекты) тоже добавлялись в ГКГ согласно заранее выбранной доле.
- 5 **Изоляция.** ГКГ исключается из всех экспериментов
- 6 **Метрики.** Те же ключевые и защитные метрики, что и в обычном А/Б-тесте
- 7 **Анализ.** В качестве драйвера экономического эффекта считается дельта ключевой метрики, усредненная по объектам рандомизации в окне W (например, 7/30 дней). Необходимо смотреть оценку и 95% ДИ.

Частые проблемы:

- 1 **Нестабильность отбора.** Главной проблемой ГКГ является сложность технической реализации и поддержание репрезентативной ГКГ на долгом горизонте. Репрезентативность легко ломают разовые сбои сплит-алгоритма или невнимательная реализация новой функциональности/модели. Типичные причины: нестабильные идентификаторы (слияние/разделение), незаметные изменения критериев включения в тест, коллизии с параллельными экспериментами.
- 2 **Нерепрезентативность ГКГ.** Слишком маленькая доля ГКГ дает высокий шум и «прыгающие» оценки, перекося по сегментам (платформа, регион, тип клиента, активность) делает непереносимой на весь портфель, со временем дрейф трафика/изменения критериев включения может приводить к тому, что группа перестает описывать текущую популяцию.

Меры по предотвращению проблем:

- Стабильный идентификатор объекта рандомизации;
- Единый источник флага принадлежности ГКГ;
- Запрет переназначений между группами;
- Четкая фиксация критериев попадания в ГКГ;
- Полное перемешивание ГКГ при любых изменениях критериев;

- Обязательные мониторинги и оповещения по здоровью ГКГ: переназначения групп, контроль слияния/разделения аккаунтов, SRM по всей популяции и по ключевым срезам.

Пересборка Глобальной Контрольной Группы:

При обнаружении перекоса в ГКГ необходимо остановить анализ, зафиксировать и устранить причину и выполнить перемешивание ГКГ. Создать новую версию ГКГ параллельно текущей, запустить в теневом режиме (без влияния на объекты воздействия) на 3–7 дней, проверить равномерность сплита по всему трафику и по ключевым срезам и проверить баланс ковариат. Переключить продукт на новую версию ГКГ и выдержать буферный период, данные переходного окна в оценку не включать (длина буферного периода должна быть не менее длины окна замера эффекта W). Далее считать экономический эффект по новой версии ГКГ.

II.II. Квази-экспериментальные методы

Если провести рандомизированный эксперимент невозможно — из-за юридических или технических ограничений, высокой стоимости или нарушения условий его применимости (SUTVA), — можно использовать альтернативные квази-экспериментальные методы. Однако стоит учитывать, что по силе причинной идентификации эффекта они значимо уступают классическим А/Б и опираются на довольно жесткие допущения о данных. В данной секции описаны только базовые принципы работы квази-экспериментальных методов, перед практическим применением стоит отдельно внимательно изучить тонкости и детали использования методов, а также учесть ограниченную область их применимости,

Switchback-тестирование

Принцип работы

Воздействие «тест/контроль» применяется не по объектам, а по временным слотам, регулярно «переключая» назначение. Подходит для систем с сильной интерференцией: ценообразование, аллокация операторов/курьеров, ранжирование, где общий рынок/совокупность ресурсов «перетекает» между группами.

Пример. Динамическое ценообразование курьеров: переключаемся по часовым блокам внутри города и сравниваем выручку на пользователя и время доставки, контролируя «час:день».

Ключевые допущения

- Временные слоты сравнимы после учета сезонности/шоков, перенос эффекта между слотами либо отсутствует, либо гасится буферным окном.
- Назначение слотов в «тест/контроль» действительно случайно (или псевдослучайно) и сбалансировано по календарным факторам.

Дизайн

- Единица рандомизации: временной слот (например, 30–60 мин) × локация/кластер.
- Расписание: случайные или чередующиеся блоки Тест/Контроль с равными экспозициями по часам/дням. Необходимо добавлять период вымывания эффектов при смене состояния, который не будет учитываться в оценке.
- Метрики: те же ключевая и защитные метрики, агрегированные окне W .

Риски

- Неверная длина слота (слишком короткий: шум/перетекания, слишком длинный: низкая мощность).
- Непокрытая сезонность (попадание в «сильные» часы только тестом).
- Переток эффекта в соседние слоты - лечится буферным окном

Propensity Score Matching (PSM)

Принцип работы

Метод позволяет сравнить тех, на кого было оказано воздействие, с очень похожими объектами, среди тех, на кого воздействия не было, при раскатке изменения сразу на всю популяцию объектов воздействия, без проведения честного А/Б-теста. С помощью МЛ-модели подбираются «близнецы» по наблюдаемым признакам: строится модель предсказывающая вероятность попасть под воздействие новой функциональности/алгоритма (propensity score). Далее объекты с фактически примененным воздействием и без сопоставляются по значению propensity score, тем самым группы выравниваются по наблюдаемым факторам. После сравниваются метрики двух групп, средняя разница и есть оценка эффекта для получивших воздействие. Метод подходит для случаев, когда рандомизация невозможна, но есть сильные пред-тестовые данные, которые описывают различия между группами, получивших и не получивших воздействие.

Пример. Запуск ИИ-ассистента продаж для B2B-клиентов: ассистент раскатывается на всех клиентов, затем подключившиеся клиенты матчатся с теми, кто не пользовался ассистентом по отрасли, размеру, прошлой выручке, и т.д. Считается дельта средней выручки по сматченным группам.

Ключевые допущения

- При условии фиксированных ковариат назначение в тест «как бы случайно».
- Есть перекрытие распределений propensity score между группами.

Дизайн

- Ковариаты: только пред-тестовые данные (история покупок/обращений, чек, гео, устройство, тариф и т.д.)
- Модель: логистическая регрессия или бустинг
- Проверки: баланс ковариат

Риски

- Ненаблюдаемые сильные ковариаты. *Пример: в пилот LLM-помощника взяли только звезд-операторов по личной оценке руководителя, о чем не знала продуктовая команда.*
- Плохое перекрытие распределений: среди получивших воздействие нет похожих объектов на не получивших. *Пример: все получившие воздействие пользователи iOS, не получившие преимущественно пользователи Android.*

Синтетический контроль (Synthetic Control)

Принцип работы

Метод позволяет оценить причинный эффект интервенции для одной или нескольких агрегированных единиц (город, линия бизнеса), когда классический А/Б невозможен. Для объекта, к которому применено воздействие строится синтетический двойник из нескольких объектов, которые не получили воздействие.

Подбирается взвешенная смесь донорских объектов, которая очень точно повторяет поведение объекта с примененным воздействием на пред-тестовых данных. Разница ключевой метрики между оригинальным объектом и синтетическим близнецом на периоде после интервенции является оценкой эффекта воздействия.

Пример: в райдтех сервисе в Казани включили новый алгоритм распределения курьеров, а в похожих городах (Нижний Новгород, Самара, Уфа) оставили старый алгоритм. Из похожих городов собрали синтетического двойника как 0.3 Новгорода + 0.5 Самары + 0.2 Уфы, который довольно точно повторял траекторию Казани до запуска изменения.

Ключевые допущения

- **Аппроксимация:** объект с воздействием в пре-периоде хорошо описывается выпуклой комбинацией объектов-доноров.
- **Отсутствие интерференции между единицами:** воздействие не искажает поведение объектов-доноров.

Риски

- Плохая подгонка на пре-периоде: синтетический близнец не похож на факт
- Слишком короткая история наблюдений перед тестом
- Сильный внешний шок на пост-периоде, который оказывает различное влияние на объекты. Например, новый закон или COVID

II.III. Альтернативное моделирование

На уровне 2 альтернативное моделирование применяется в том же виде, что описано в уровне 1 (AS IS / TO BE). Отличие — более длинный горизонт оценки и необходимость учитывать косвенные эффекты на смежные процессы и продукты при портфельной оценке.

См. описание метода в уровне 1.

II.IV. Отраслевая специфика

На уровне 2 отрасли расходятся по тому, какой продвинутый метод становится основным: усиленный A/B и Holdout — в отраслях с массовыми экспериментами, switchback — при сильной интерференции, Synthetic Control и PSM — при оценке на агрегированных единицах и в B2B.

Финансовая индустрия

Метод: Усиленный A/B (стратификация, CUPED); вечный A/B с глобальной контрольной группой (Holdout); Propensity Score Matching.

Почему: У зрелых моделей ожидаемый эффект мал, поэтому нужна повышенная чувствительность — её дают стратификация и CUPED, сокращая необходимую выборку. Для крупных постоянно работающих моделей (скоринг, рекомендации) ставят глобальную контрольную группу: она отслеживает дрейф и честно считает интегральный эффект. В B2B-сегменте, где клиентов мало и раскатка идёт сразу на всех, применяют Propensity Score Matching.

Ритейл и e-com

Метод: Switchback; вечный A/B с глобальной контрольной группой (Holdout) на уровне продукта; Synthetic Control.

Почему: Курьеры — общий ресурс: воздействие на одну группу заказов меняет их доступность для другой, и обычный А/Б нарушает SUTVA. Для ценообразования и аллокации поэтому используют switchback — переключение тест/контроль по временным слотам.

Сумма эффектов отдельных А/Б по фичам не сходится с общим приростом товарооборота, поэтому ставят постоянную контрольную группу на уровне всего продукта; её выделение и удержание организационно сложно и требует отдельной договорённости. Для оценки на уровне города или сети дарксторов, где нет сравнимой пары, собирают синтетического двойника (Synthetic Control).

Транспорт и логистика

Метод: Switchback; Synthetic Control.

Почему: Ценообразование и распределение ресурсов — процессы с сильной интерференцией: А/Б по цене может дать даже неверный знак эффекта, потому что участники рынка влияют друг на друга. Поэтому тест/контроль переключают по временным слотам (switchback). Для оценки на уровне маршрута или региона строят синтетического двойника из похожих маршрутов и регионов (Synthetic Control).

АПК, энергетика и промышленность

Метод: Synthetic Control; Propensity Score Matching.

Почему: Там, где есть несколько сопоставимых агрегированных единиц (линия, цех, актив, месторождение), единицу с ИИ сравнивают с синтетическим двойником, собранным из остальных (Synthetic Control). Где единицы можно сопоставить по наблюдаемым признакам — применяют Propensity Score Matching. Эффекты «непрожитой реальности» и ускорения инвест-проектов на этом уровне ещё трудноизмеримы и переходят на уровень 3.

ИКТ

Метод: Усиленный А/Б; вечный А/Б с глобальной контрольной группой (Holdout); Propensity Score Matching.

Почему: Для массовых процессов (удержание абонентов, маркетинговые коммуникации) применяют усиленный А/Б и глобальную контрольную группу. В B2B-сегменте, где крупных клиентов немного и раскатка идёт сразу на всех, — Propensity Score Matching. Из-за мультисегментности (B2C / B2B / HR) метод в каждом сегменте выбирается отдельно.

АФИИНА III

III.I. Альтернативное моделирование на исторических и рыночных данных

На уровне АФИИНА III А/Б-тест и квази-эксперименты, как правило, неприменимы: горизонт оценки длинный (3–5+ лет), эффект системный и не локализован в конкретной группе пользователей или процессе, а контрольной группы «без ИИ» на уровне отрасли не существует. Меняется и предмет оценки — от инкремента отдельной инициативы к вкладу ИИ в производительность и конкурентоспособность компании или крупного бизнес-направления. Основным методом становится альтернативное моделирование — тот же метод, что описан на уровнях 1 и 2, но применённый на исторических и рыночных данных и на длинном горизонте.

Важно удержать природу метода из блока I.II: альтернативное моделирование — это приближённая, экспертная оценка экономического эффекта на основе бизнес-моделирования. Оно строит модель «как должно работать» и рассчитывает порядок величины эффекта при заданных допущениях. Это не квази-

эксперимент и не доказательство причинного влияния; его сила — в прозрачности механики, а не в статистической строгости причинности.

Суть метода: AS IS / TO BE через цепочку драйверов

Эффект в альтернативном моделировании строится не вычитанием двух макро-траекторий, а моделированием механизма эффекта — явной цепочки «применение ИИ → изменение метрики → деньги», наложенной на базовый сценарий «без ИИ».

Метод опирается на два элемента:

- **Базовый сценарий (AS IS, без ИИ)** — как вёл бы себя показатель без внедрения ИИ: тренд, сезонность, макроэкономический и отраслевой фон. Это точка отсчёта.
- **Целевой сценарий (TO BE, эффект)** — как меняется показатель под воздействием ИИ, рассчитанный через цепочку драйверов и переведённый в деньги.

Итоговый эффект — разница целевого и базового сценариев в сценарной вилке Base / Low / High. На уровне АФИИНА III отличие только в масштабе и данных: и базовый сценарий, и коэффициенты драйверов калибруются на длинных исторических рядах компании и внешних рыночных данных.

Данные

Модель строится на двух типах данных:

- **Исторические данные компании** — длинные временные ряды ключевых показателей (выручка, затраты, качество портфеля, операционные метрики) на уровне всей компании или крупного бизнес-направления. На них калибруются базовый сценарий и коэффициенты драйверов.
- **Внешние рыночные данные** — отраслевые индикаторы, макроэкономические показатели, данные сопоставимых компаний. Используются как контрольные факторы и ковариаты: позволяют отделить эффект ИИ от общей рыночной и макроэкономической динамики.

Особенности и ограничения уровня

Три ограничения определяют границы применимости любого инструмента на этом уровне:

- **Отсутствие контрольной группы.** При массовом внедрении «отрасль без ИИ» не существует — сравнивать не с кем, поэтому базовый сценарий строится модельно, а не наблюдается.
- **Интерференция и равновесные эффекты.** Внедрение ИИ у одного игрока меняет цены и доли рынка у остальных; классическое допущение независимости наблюдений нарушается, а перетоки и конкурентная реакция являются частью измеряемого эффекта.
- **Длинный горизонт и накопление факторов.** За 5+ лет накапливаются макроэкономические, регуляторные и структурные сдвиги, поэтому отделение эффекта ИИ от прочих факторов становится критичным.

Прямое следствие: результат уровня АФИИНА III подаётся диапазоном по сценариям при документированных допущениях и подтверждается триангуляцией — совпадение оценок из независимых подходов повышает доверие, расхождение является сигналом к разбору, а не поводом выбрать удобную цифру.

Как повысить корректность альтернативного моделирования

Главная слабость метода — не базовый сценарий, а **коэффициенты драйверов** (условные «-X% времени обработки», «+Y% конверсии»). Обычно это экспертные числа, и именно здесь оценка проседает. Поэтому

наибольший вклад в корректность дают инструменты, переводящие драйверы с экспертной оценки на историческую статистику. Ниже — по звеньям цепочки.

1. Базовый сценарий «без ИИ». Вместо линии «на глаз» — статистическая декомпозиция ряда на тренд, сезонность и остаток (STL, ARIMA/SARIMA, ETS/Holt-Winters) либо регрессия базового сценария на внешние факторы (макро, цены, сезонные драйверы). Это даёт воспроизводимую точку отсчёта.

02 Коэффициенты драйверов — основной резерв корректности.

- **Эконометрическая оценка эластичностей** на исторических данных: как целевая метрика реально реагирует на драйвер, с контролем прочих факторов — вместо экспертного коэффициента.
- **Перенос коэффициента из пилота или А/Б-теста** с его доверительным интервалом, а не точечным значением: подтверждённая на уровнях 1–2 дельта становится обоснованным входом в модель.
- **Инструментальные переменные и контроль конфаундеров** — когда драйвер эндогенен (ИИ внедряли там, где метрика и так росла), чтобы коэффициент не вобрал постороннее влияние.
- **Регуляризация и отбор факторов**, когда драйверов много и они коррелированы между собой.

3. Проброс неопределённости через цепочку. Монте-Карло и сценарии Base / Low / High из блока I.II привязываются не только к сценариям выручки, но и к **распределениям коэффициентов драйверов**; для доверительных интервалов оценённых эластичностей применяется bootstrap. Итог — честный диапазон, а не одна цифра.

04 Валидация модели.

- **Бэкtest базового сценария** на отложенном доинтервенционном окне (train/test на исторических рядах): модель должна воспроизводить известное прошлое, прежде чем ей доверять на контрфакте.
- **Проверка остатков** — автокорреляция, гетероскедастичность, стационарность; иначе доверительные интервалы занижаются.
- **Анализ чувствительности** к ключевым допущениям и **сверка с PnL и бенчмарками** (в I.II это «валидация здравым смыслом» — здесь она формализуется).

Эти инструменты не меняют природу метода — он остаётся приближённым и на допущениях, — но убирают его главную слабость: произвольность коэффициентов драйверов и ненадёжность базового сценария.

Контрфактическая проверка на агрегированных рядах (условное дополнение)

Отдельно от структурного моделирования результат можно перепроверить эконометрическими контрфактическими методами. Это не ядро уровня, а независимая сверка: они не моделируют механизм драйверов, а статистически реконструируют агрегированный контрфакт «без ИИ» и приписывают разрыв эффекту (логика «факт – контрфакт»).

- **Структурные модели временных рядов (BSTS / causal impact)** — наиболее реализуемый вариант. По доинтервенционному ряду и коррелирующим контрольным рядам строится вероятностный прогноз контрфакта с доверительным интервалом; эффект — разница факта и прогноза. Метод имеет готовый инструментарий и применяется в индустрии для оценки запусков.
- **Прерванный временной ряд (ITS)** — более простой и грубый вариант: держится на допущении, что доинтервенционный тренд сохранился бы. Годится как первичная проверка, а не как основная оценка.
- **Векторные авторегрессии (VAR/VECM)** — опциональный продвинутый инструмент для взаимосвязанных отраслевых рядов; на практике редко реализуем как управленческий, так как требует измеримого «шока

внедрения» и длинных стабильных рядов.

Условия реализуемости. Контрфактическая проверка выполнима на уровне компании или крупного бизнес-направления при одновременном выполнении условий:

- датируемый момент раскатки (при плавной, размазанной на годы диффузии ИИ разрыв смазывается и методы теряют мощность);
- длинный чистый ряд до внедрения — 3–5 лет месячных или квартальных данных;
- контрольные ряды, не затронутые ИИ (для BSTS): на уровне компании рыночные индексы и макро годятся, поскольку одна компания обычно не двигает весь рынок;
- отсутствие крупного постороннего шока в пост-периоде, иначе разрыв ошибочно припишется ИИ.

На уровне отрасли или экономики контрфактическая проверка нереализуема: чистых контрольных рядов нет (все внедрили ИИ), а адопция размазана во времени. Поэтому метод применяется как проверка на уровне компании, а не как универсальный инструмент системного масштаба.

Принцип триангуляции

Ни один метод уровня АФИИНА III не даёт строгой причинности. Поэтому итоговая оценка берётся через схождение двух логик: структурного альтернативного моделирования (снизу вверх, через драйверы) и контрфактической проверки (сверху вниз, где она реализуема). Совпадение оценок повышает доверие; расхождение — сигнал к разбору допущений, а не повод выбрать удобное значение.

Ключевые требования к модели

- **Длинный исторический ряд** — не менее 3–5 лет до внедрения для надёжной калибровки базового сценария и коэффициентов драйверов, устранения случайных флуктуаций.
- **Явное разделение эффекта ИИ и прочих факторов** — макроэкономика, регуляторные изменения, конкурентная среда и органический рост выделяются отдельно и либо исключаются из модели, либо учитываются как ковариаты.
- **Обоснование коэффициентов драйверов** — каждый коэффициент цепочки имеет источник: эконометрическая оценка, перенос из пилота с доверительным интервалом или явно помеченное экспертное допущение.
- **Проверка допущений** — бэк-тест базового сценария, диагностика остатков; для контрфактической проверки — устойчивость тренда (ITS), качество подгонки на пред-периоде (BSTS), валидность инструмента (IV).
- **Чувствительностный анализ** — ключевые параметры проверяются в сценариях Base / Low / High; итог представляется диапазоном, а не точечной оценкой.
- **Триангуляция и документирование** — где возможно, оценка подтверждается независимым методом; все ключевые допущения документируются. Без независимой валидации модель системного уровня не используется для внешнего репортинга.

Ограничения метода

Альтернативное моделирование на уровне АФИИНА III по природе своей менее точно, чем А/Б-тест или квази-эксперимент. Оно не устанавливает причинно-следственную связь в строгом статистическом смысле, а строит обоснованную оценку эффекта при явно сформулированных допущениях. На системном горизонте невозможно ни поставить эксперимент, ни полностью отделить ИИ от одновременно идущих изменений экономики. Поэтому ценность оценки определяется не точностью одной цифры, а прозрачностью допущений,

честностью диапазона и согласованностью результатов, полученных разными путями: читатель должен иметь возможность воспроизвести логику расчёта и оспорить конкретное допущение, а не всю модель целиком.

Финансовая отрасль

01 Кампании продаж и модели приоритизации клиентов на основе ИИ

ФИНАНСОВЫЙ ЭФФЕКТ

182 млн. руб.

СБЕР · за 2023 г.

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Продуктовая конверсия

СЕГМЕНТ КЛИЕНТОВ

B2B

ТИП ИЗМЕНЕНИЯ

Автоматизация

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Увеличить процент совершения сделок по продуктам, с которыми наиболее глубоко взаимодействовал клиент, перевести "брошенные" заявки в заключенные сделки.

ТИП ЭФФЕКТА

Эффект на доход

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

Рандомизированный А/Б тест

ОПИСАНИЕ ТЕХНОЛОГИИ

Классификационные модели (CatBoost, и т.п.)

Описание

Цель модели: Ранжировать продуктовые предложения по итогу сессии взаимодействия клиента с продуктами на сайте sberbank.ru / в личном кабинете Сбербизнес.

Пример расчета

КАЛЬКУЛЯТОР

Финансовый Эффект_T = $\sum \text{ДельтаКонверсии}_i * \text{БазаЛидов}_j * \text{PL}_{ij}$, где T – месяц, к которому проводится оценка от инициативы; N – количество продуктов сплита; ДельтаКонверсии_i – это дельта конверсии i-ого продукта сплита, полученная по пилоту; БазаЛидов_j – это база лидов в месяц j; PL_{ij} – это значения PL i-ого продукта в j-ый месяц жизни этого продукта. При этом, поскольку пилот

проводился только 3 месяца (нет постоянного контроля), на значение ДельтаКонверсии_i выставляется мониторинг, корректирующий значения конверсии в случае необходимости.

02 Чат-бот клиентского обслуживания

ФИНАНСОВЫЙ ЭФФЕКТ

433 млн. руб.

Т БАНК · в мес.

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Время операторов поддержки

ТИП ЭФФЕКТА

Расходы

СЕГМЕНТ КЛИЕНТОВ

B2C

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Вся активная клиентская база Т-банка

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

А/Б-тест с вечным контролем (holdout)

ТИП ИЗМЕНЕНИЯ

Автоматизация

ОПИСАНИЕ ТЕХНОЛОГИИ

NLP

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Чат-бот клиентского обслуживания: распознает намерения клиента с помощью ML и дает релевантные ответы без участия человека, может предоставить информацию или решить запрос прямо в чате.

Описание

Расчет экономического эффекта осуществляется с помощью АБ-теста с вечным контролем относительно альтернативного сценария работы поддержки без чат-бота.

Дизайн эксперимента:

Сплит по клиентам при обращении в чат поддержки.

Контроль (1%) - боты отключены, клиент попадает сразу на операторов;

Тест (99%) - сначала помогает бот

Пример расчета

КАЛЬКУЛЯТОР

$\Delta t = (\text{время операторов за месяц в контроле} / \text{кол-во клиентов в контроле} - \text{время операторов за месяц в тесте} / \text{кол-во клиентов в тесте}) * \text{кол-во клиентов в тесте} = \text{время операторов за месяц в контроле} * 99 - \text{время операторов за месяц в тесте} = 2,11 \text{ млрд. сек.}$

$\text{Экономия} = \Delta t * \text{стоимость секунды оператора} = 2,11 \text{ млрд. сек.} * 0.205 \text{ руб/с} = 433 \text{ млн. руб. в мес.}$

03

Модель определения дохода клиента в целях расчета Показателя Долговой Нагрузки

ФИНАНСОВЫЙ ЭФФЕКТ

23%

Альфа Банк · ожидаемая доходность

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

eROE, потребление капитала

ТИП ЭФФЕКТА

Эффект на доход

СЕГМЕНТ КЛИЕНТОВ

B2C

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

A/B тестирование

ТИП ИЗМЕНЕНИЯ

Автоматизация

ОПИСАНИЕ ТЕХНОЛОГИИ

ML (градиентный бустинг – утверждено с ЦБ)

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Десятки миллионов клиентов (клиентская база АЛЬФА-БАНКА)

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Расчет Показателя Долговой Нагрузки при помощи модельного подхода к оценке дохода клиента

Описание

Получили премию [GIA 2024](#) – Ежегодная премия в области корпоративных инноваций

В 2023 г. АО «АЛЬФА-БАНК» в тройке первых получил разрешение от Банка России о допустимости применения собственных моделей для расчета Показателя долговой нагрузки, в 2025 г. получил разрешение на обновленную модель. Благодаря разработанным моделям АО «АЛЬФА-БАНК» более персонализировано предлагает свои продукты, более точно просчитывает риски невозврата, при этом улучшает клиентский путь.

Пример расчета

КАЛЬКУЛЯТОР

$$\text{eROE} = \text{ожидаемая чистая прибыль} / \text{потребление капитала} = \text{ожидаемая чистая прибыль} / (\text{RWA} * \text{CAR})$$
$$\text{eROE}_{\text{без_модели}} = 10\% / (2.5 * 11.5\%) = 35\%$$
$$\text{eROE}_{\text{с_моделью}} = 10\% / (1.5 * 11.5\%) = 58\%$$
$$\Delta \text{eROE} = \text{eROE}_{\text{с_моделью}} - \text{eROE}_{\text{без_модели}} = 23\%$$

eROE – ожидаемая доходность

RWA – активы, взвешенные по уровню риска

CAR – норматив достаточности капитала

Чистая прибыль ↑, объемы бизнеса ↑, удовлетворенность клиентов ↑, потребление капитала ↓.

04 Прогнозирование метрик регулярной отчетности с помощью AutoML

ФИНАНСОВЫЙ ЭФФЕКТ

10,67 млрд. руб/год

MOEX GROUP

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

МАРЕ, трудозатраты

СЕГМЕНТ КЛИЕНТОВ

Внутренний клиент

ТИП ИЗМЕНЕНИЯ

Автоматизация

ТИП ЭФФЕКТА

Потенциальные расходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

Моделирование

ОПИСАНИЕ ТЕХНОЛОГИИ

AutoML

Описание

Для повышения качества прогнозирования в регулярной отчетности, необходимо замещение классических статистических методов прогнозирования (в т.ч. run-rate) на прогнозирование методами машинного обучения, что улучшает точность прогноза по каждой метрике от 5% до 60 раз. Т.к. прогнозировать нужно более 1500 метрик, то при еженедельной периодичности прогноза, при регулярном re-feet моделей, требуется дополнительное привлечение огромного штата датасаентистов. Использование AutoML решает эту проблему. При подаче данных происходит формирование объясняющих факторов, отбор метода ML из 378 (учитывая стекинг) и автоматический фиче.инжиниринг (отбор факторов). На выходе бизнес-аналитик автоматически получает модель ML и прогноз. По сути инструмент автоматизирует работу датасаентистов, которые являются достаточно дорогими сотрудниками.

Пример расчета

— КАЛЬКУЛЯТОР

Фин.эффект = 1500 моделей X 5 чел.-дней X 52 расчётов/год = 390 тыс.чел.-дней./год = 1611 ПШЕ / год = 10,67 млрд.руб./год
Трудозатраты на реализацию = 80 чел.-дней = 2,2 млн.руб.; Трудозатраты на поддержку = 40 чел.-дней/год = 1,1 млн.руб./год.; ROI* = $(10,67 \text{ млрд.руб./год} - 3 \text{ года} - 1,1 \text{ млн.руб./год} \cdot 3 \text{ года} - 2,2 \text{ млн.руб.}) / (1,1 \text{ млн.руб./год} \cdot 3 \text{ года} + 2,2 \text{ млн.руб.}) = 581900\%$

*при расчёте на 3 года.

05 AI-агент по формированию one-page-слайда по презентации

ФИНАНСОВЫЙ ЭФФЕКТ

50%

MOEX GROUP · ROI

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Трудозатраты

СЕГМЕНТ КЛИЕНТОВ

Внутренний клиент

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Блок финансы

ТИП ЭФФЕКТА

Потенциальные расходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

Моделирование

ОПИСАНИЕ ТЕХНОЛОГИИ

LLM, AI-агент

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Формирование one-page

Описание

Бизнес аналитикам приходится периодически формировать one-page слайд на основе большой презентации. Это долго и трудозатратно. Использование AI-агента для формирования one-page сокращает трудозатраты с многих часов до нескольких секунд. Агент принимает на вход презентацию pptx и преобразует элементы презентации (текст, графики, таблицы, диаграммы, картинки) в формат удобный для LLM. Далее LLM делает саммари материала и выдает результат в определенном xml-подобном формате. Далее агент через специализированные python библиотеки формирует из этого «xml» слайд презентации pptx.

Пример расчета

— КАЛЬКУЛЯТОР

Фин.эффект = 1 рабочий день * 50 расчётов/год = 50 дней./год = 0,2 ПШЕ / год
Трудозатраты на реализацию = 40 рабочих дней; Трудозатраты на поддержку = 20 рабочих дней;
 $ROI^* = (503 - 203 - 40) / (20 * 3 + 40) = 50\%$.



Скоринг для исходящих коммуникаций с клиентами банка в Контактном центре

ФИНАНСОВЫЙ ЭФФЕКТ

+104 млн руб.

ПСБ · за 2025 г. · (ожидаемый эффект на момент заполнения информации)

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

CR/NPV/PnL

СЕГМЕНТ КЛИЕНТОВ

B2C (клиенты банка)

ТИП ИЗМЕНЕНИЯ

Оптимизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Рекомендации по продуктам Банка в КЦ

ТИП ЭФФЕКТА

Доходы/оптимизация расходов

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

A/B тестирование: ML против случайной выборки

ОПИСАНИЕ ТЕХНОЛОГИИ

ML (модели классификации)

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Персональные рекомендации, оптимизация конверсии по продажам, повышение эффективности КЦ

Описание

Разработаны ML-модели оценки вероятности отклика на продукты Банка после коммуникации. Клиенты с высокой вероятностью отклика на продукт попадали в списки для исходящих звонков.

Пример расчета

— КАЛЬКУЛЯТОР

$$\text{ФинЭффект} = \sum (\text{СегментКлиентов_ML} \times (\text{CR_ML} - \text{CR_control}) * \text{PNL_li} - \text{E_i}) - \text{E_ML}$$

CR_ML - отклик в группе отобранной по модели, CR_control - отклик в случайно отобранной группе
PNL_{1i} – PnL за один год по продукту i
E_i - затраты на проведение коммуникаций по продукту i, включая ЦГ и КГ
E_ML - затраты на инфраструктуру и ФОТ для разработки ML модели

07

Создание AI-аватара для сотрудников поддержки

ФИНАНСОВЫЙ ЭФФЕКТ

+9%

Банк из ТОП-15 · Рост кол-ва положительных оценок на FTE

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

CSAT, eNPS

СЕГМЕНТ КЛИЕНТОВ

B2B / B2C / B2E

ТИП ИЗМЕНЕНИЯ

Реальные фото → Генеративные

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

500 операторов

ТИП ЭФФЕКТА

Расходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

A/B тестирование, месяц

ОПИСАНИЕ ТЕХНОЛОГИИ

GenAI (Диффузии) — Remini и Magic Eraser

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Создание однотипных аватарок для специалистов поддержки с помощью GenAI-сервисов

Описание

Мы разработали решение — ИИ-аватарки для сотрудников поддержки. Ранее каждый выбирал фото самостоятельно, что снижало единый стиль и доверие. Новая технология создаёт уникальные аватарки, решая проблему отсутствия фото и улучшая коммуникацию с клиентами. Кроме того, аватарки защищают сотрудников, исключая возможность поиска их в соцсетях недобросовестными клиентами. Внедрение повысило вовлечённость сотрудников и улучшило показатели удовлетворённости клиентов (CSAT). Проект объединяет инновации и заботу о клиентах и сотрудниках, повышая качество обслуживания и лояльность.

Пример расчета

КАЛЬКУЛЯТОР

Рост положительных оценок = (Новые положительные оценки за месяц после — Исходные положительные оценки за месяц до) / Исходные хорошие оценки × 100%

Исходные оценки: 6,400 Новые оценки: 6,976 Прирост = $(6,976 - 6,400) / 6,400 \times 100\% = 9\%$



Чат и голосовой бот клиентского обслуживания

ФИНАНСОВЫЙ ЭФФЕКТ

410 млн. руб.

Банк из ТОП-15 в год

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

PnL

СЕГМЕНТ КЛИЕНТОВ

B2C

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Активная клиентская база физических лиц банка

ТИП ЭФФЕКТА

Расходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

Серия A/B-тестов без вечного контроля

ОПИСАНИЕ ТЕХНОЛОГИИ

NLP

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Цель — сократить совокупное время операторов, сохранив или улучшив удовлетворённость клиентов, тем самым снизив прямые операционные расходы.

Описание

Чат- и голосовой бот: распознаёт намерение клиента (NLP), с помощью большой языковой модели выбирает наилучший ответ в рамках темы полностью закрывает запрос без участия человека. Частично автоматизированные диалоги (когда оператор подключается) не включаются в дальнейший расчёт эффекта.

Пример расчета

КАЛЬКУЛЯТОР

Доля полностью автоматизированных диалогов: $p = N_{\text{auto}} / N_{\text{total}}$

Сэкономленное время операторов: $\Delta T = p \times N_{\text{total}} \times \text{top}$

Экономия в деньгах: $\text{Экономия} = \Delta T \times C_{\text{сек}}$

N_auto – диалоги, полностью закрыты ботом
N_total – все диалоги за период
top – среднее время оператора на 1 диалог, сек
C_сек - себестоимость секунды работы оператора, руб
Экономия ~ **410 млн руб.** год.



Модель привлечения BARCELONA (продукт Премиум)

ФИНАНСОВЫЙ ЭФФЕКТ

+150 млн. руб.

ГАЗПРОМБАНК

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

CR, CTR

СЕГМЕНТ КЛИЕНТОВ

B2C

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

CRM система розничного бизнеса

ТИП ЭФФЕКТА

Эффект на доход

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

A/B тестирование

ОПИСАНИЕ ТЕХНОЛОГИИ

GBM модель

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Привлечение новых клиентов на пакет услуг Премиум

Описание

Модель склонности к оформлению пакета услуг Премиум на основе истории откликов клиентов.

Пример расчета

КАЛЬКУЛЯТОР

Кол-во привлеченных клиентов * LTV клиента = Итоговый эффект. image-4.jpg



Аналитический сервис по автоматизации процесса оценки и переоценки отдельных видов профильных залогов юридических лиц

(БОАИ - база отраслевой экономической информации)

ФИНАНСОВЫЙ ЭФФЕКТ

+472,6 млн руб.

РСХБ · ожидаемый доход к 2029 г.

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

коэффициент покрытия объектов типового залога

СЕГМЕНТ КЛИЕНТОВ

B2B

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Внутреннее подразделение Банка, ЮЛ

ТИП ЭФФЕКТА

Доходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

A/B-тестирование

ОПИСАНИЕ ТЕХНОЛОГИИ

ML (градиентный бустинг и линейные модели)

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Оценка залоговой стоимости имущества ЮЛ

Описание

Автоматизация процесса оценки и переоценки отдельных видов профильных залогов клиентов-юридических лиц за счет внедрения аналитического сервиса, использующего технологии ИИ

Цель реализации:

- Ускорение процесса определения стоимости отдельных видов профильного залогового обеспечения до онлайн оценки
- Сокращение прямых затрат Банка и повышение объективности оценки отдельных видов профильного залогового обеспечения
- Обеспечение возможности мониторинга стоимости активов на портфельном уровне

Пример расчета

— КАЛЬКУЛЯТОР

Применимость (покрытие) сервиса по оценке залогов – 50% номенклатуры. Вычисляется как отношение числа объектов залога, по которому сервис обеспечил переоценку в границах доверительного диапазона (+/-10% от предыдущей оценки, проведенной экспертом Банка), к общему числу объектов залога тех видов, к которым применим сервис (транспорта, техника, прицепное и навесное оборудование)

Ожидаемый доход к 2029 г./млн руб. = ежеквартальное сокращение расходов на содержание персонала (заработная плата и страховые взносы (ОМС, ОПС, ОСС) + выплаты при увольнении) X 4 квартала X количество лет реализации бизнес-кейса $(19,6928+0) \times 4 \times 6 = 472,62$.



Время на поиск ответов по бизнес Википедии и продуктовому каталогу для сотрудников колл центра (КЦ)

ФИНАНСОВЫЙ ЭФФЕКТ

~490 млн. руб.

ВТБ · на горизонте 4 лет

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Average Usefulness (AU)

СЕГМЕНТ КЛИЕНТОВ

Retail

ТИП ИЗМЕНЕНИЯ

Автоматизация

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Персональные рекомендации

ТИП ЭФФЕКТА

Эффект на расходы

МЕТОД ОЦЕНКИ ФИН ЭФФЕКТА

Измерения

ОПИСАНИЕ ТЕХНОЛОГИИ

LLM (RAG сервис)

Периметр внедрения кейса:
Сотрудники КЦ (> 100 чел)

Описание

Оптимизация трудозатрат от сокращения времени на поиск ответов сотрудниками КЦ

Пример расчета

КАЛЬКУЛЯТОР

$$E = \sum_i (N(i) (AT_Control - T1_Pilot) CostPer1min(i))$$

N(i) Количество звонков в КЦ, обрабатываемых операторами (в пилотной группе, по годам i)
AT_Control Среднее время телефонного диалога до реализации проекта (на основании данных журналирования)
T1_Pilot Среднее время телефонного диалога после реализации проекта (на основании расчетов SLA сервиса LLM+RAG)
CostPer1min(i) Стоимость 1 минуты диалога Оператора КЦ (с учетом прогноза роста по годам, i).

Ритейл и e-com

Скоро тут появятся кейсы

01 Видеоаналитика с применением ИИ при обслуживании инфраструктурных объектов

ФИНАНСОВЫЙ ЭФФЕКТ

+50%

оптимизация временных ресурсов

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Операционная эффективность

СЕГМЕНТ КЛИЕНТОВ

B2O (операторы)

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ

40 000 базовых станций сотовой сети

ТИП ЭФФЕКТА

Снижение расходов

МЕТОД ОЦЕНКИ ФИНЭФФЕКТА

Оценка экономии операционных затрат (Cost Savings)

ТЕХНОЛОГИИ

CV компьютерное зрение

ОПИСАНИЕ БИЗНЕС-ЗАДАЧИ

Сокращение трудозатрат на обслуживание базовых станций сотовой сети

Описание

Решение на основе компьютерного зрения автоматизирует инвентаризацию и техническое обслуживание базовых станций сотовой сети. Технические бригады и / или БПЛА выполняют фото и видеосъемку оборудования, которая автоматически обрабатывается CV моделью: система распознает тип оборудования и выявляет отклонения. Результаты формируются в инвентаризационные ведомости. Внедрение позволяет сократить трудозатраты на обслуживание одной станции на 20% и снизить кол-во ошибок.

Пример расчета

— КАЛЬКУЛЯТОР

Сокращение времени инвентаризации на 50%: $13\,300 \text{ станций} * 15 \text{ мин} * 50\% * 700 \text{ руб} / \text{час} = 1,3 \text{ млн руб} / \text{год}$

Сокращение времени тех обслуживания на 18%: $40\,000\text{ ТО} * 60\text{ мин} * 18\% * 700\text{ руб} / \text{час} = 5\text{ млн руб} / \text{год}$

Отсутствие трудозатрат на обработку фотоотчетов: $40\,000\text{ ТО} * 10,5\text{ мин} * 700 = 4,9\text{ млн руб} / \text{год}$

Сокращение использования автовышки 5%: $12\,000\text{ станций} * 5\% * 3500\text{ руб/час} * 4\text{ часа} = 8,4\text{ млн руб} / \text{год}$

02 AI Engineer для автоматизации инцидент-менеджмента сети

ФИНАНСОВЫЙ ЭФФЕКТ

до 95% точность работы ИИ-инженера

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Скорость обработки инцидентов, уровень автоматизации, FTE

СЕГМЕНТ

Внутренние операции

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ КЕЙСА

Пилот проходил в нескольких регионах с декабря 2025 по февраль 2026 года. Весной 2026 года решение масштабировано на всю страну

ТИП ЭФФЕКТА

Снижение расходов, связанных с обработкой инцидентов

МЕТОД ИЗМЕРЕНИЯ ФИНАНСОВОГО ЭФФЕКТА

Повышение производительности труда

ОПИСАНИЕ ТЕХНОЛОГИИ

GenAI + ML + анализ телеметрии и логов

ОПИСАНИЕ БИЗНЕС ЗАДАЧИ

Автоматизация диагностики, маршрутизации и обработки сетевых инцидентов

Описание

AI Engineer анализирует телеметрию оборудования, журналы событий, топологию сети и исторические инциденты. На основе этих данных система автоматически определяет вероятную причину неисправности, выполняет команды на оборудовании, в том числе рестарты, и при необходимости маршрутизирует на необходимую группу и фиксирует полную диагностику в тикете. Решение позволяет сократить объем ручных операций первой линии поддержки, ускорить обработку инцидентов и повысить уровень автономности процессов мониторинга сети.

Пример расчета

Итоги к 3Q26:

ИИ-инженер обрабатывает 40% от полного потока инцидентов и самостоятельно принимает меры для решения проблемы примерно в 15% случаев (от полного потока) — без подключения первой линии сопровождения

320 000 инцидентов обработано к середине июля 2026 года

С помощью ИИ-агента время взятия инцидента в работу сократилось в 7 раз

Промышленность, энергетика и АПК

ИИ-агент подбора материалов из запасов компании

Сценарий GenAI для снижения ЧОК и предотвращения избыточной закупки

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

Снижение ЧОК, рост EBITDA

СЕГМЕНТ КЛИЕНТОВ

B2B

ТИП ИЗМЕНЕНИЯ

Автоматизация

ПЕРИМЕТР ВНЕДРЕНИЯ

Заполярный и Кольский филиалы Компании

ТИП ЭФФЕКТА

Снижение расходов

МЕТОД ОЦЕНКИ ФИНЭФФЕКТА

Комбинированный

ТЕХНОЛОГИИ

GenAI

ОПИСАНИЕ БИЗНЕС-ЗАДАЧИ

Увеличить объем вовлечения материалов из запасов Компании, экономия на закупках

Описание

Система с искусственным интеллектом, производящая подбор замен и аналогов строительных материалов и оборудования из наличия со складов Компании. Предотвращает избыточную закупку.

Система адаптирована под несколько ключевых сценариев бизнес-процесса и имеет 3 режима работы, как подбор МТР по ведомости материалов, так и массовая обработка загружаемых файлов размером несколько тысяч позиций, а также, функционал свободного поиска через поисковую строку.

Учитывает адресность, срок нахождения на складе, а также стоимость логистики при перемещении со склада на объект.

Пример расчета

КАЛЬКУЛЯТОР

$$FCFF = EBITDA \times (1 - t) + D\&A \times t - CapEx - \Delta NWC$$

Эффект на FCFF за счет снижения NWC

Если компания cash negative – доп.эффект на процентные платежи

Транспорт и логистика

Единая платформа коммуникаций и чат-бот клиентской поддержки

ФИНАНСОВЫЙ ЭФФЕКТ

-155 млн руб

снижение расходов на контакт-центр

Кейс автоматизации внешних и внутренних коммуникаций

Основные факты

КЛЮЧЕВЫЕ БИЗНЕС-МЕТРИКИ

CSI, NPS

СЕГМЕНТ КЛИЕНТОВ

B2C, B2E

ТИП ИЗМЕНЕНИЯ

Аугментация

ПЕРИМЕТР ВНЕДРЕНИЯ

8 млн/год коммуникаций с пассажирами

ТИП ЭФФЕКТА

Снижение расходов

МЕТОД ОЦЕНКИ ФИНЭФФЕКТА

Альтернативное моделирование

ТЕХНОЛОГИИ

NLU, LLM (GigaChat), RAG, ASR, TTS

ОПИСАНИЕ БИЗНЕС-ЗАДАЧИ

Чат-бот, голосовой ассистент, речевая аналитика

Описание

ИИ-система автоматизации внешних и внутренних коммуникаций реализует семантический анализ запросов (NLU), подготовку ответов при помощи LLM и корпоративной базы знаний (RAG-архитектура). В процессе взаимодействия используются исторические данные: применяется персонализированный подход и формируется уникальный опыт.

Пример расчета

КАЛЬКУЛЯТОР

a — стоимость 1 сек. обслуживания, руб

b — количество обращений в год, шт

t — среднее время обслуживания без внедрения проекта, сек

Δt — сокращение среднего времени обслуживания за счёт ИИ, сек

$$ЭЭ = t \times b \times a - (t - \Delta t) \times b \times a$$

АФИИНА

АФИИНА — открытая база знаний.

УЧАСТНИКИ ПРОЕКТА



НОРНИКЕЛЬ



Ростелеком



АЭРОФЛОТ
Российские авиалинии



Билайн

купер



АЛИАНС
В СФЕРЕ
ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА