НОВАЯ ЭРА НАУКИ И ЗНАНИЙ
МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЖУРНАЛ
global-science-review.com
editor@global-science-review.com

ЭКОНОМИКА ОШИБОК ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: СКОЛЬКО СТОИТ НЕПРАВИЛЬНОЕ РЕШЕНИЕ АЛГОРИТМА?

Адизова Мадина Абдулазиз кизи

“студент 4 курса”

Факультет Экономики,

Кафедра Экономика транспорта, http://orcid.org/ORCID:0009-0007-9785-1117

Ташкентский Государственный Транспортный Университет , Ташкент, Узбекистан.

*e-mail:**madinaadizova59@gmail.com***

Аннотация. Цель. Разработать подход к экономической оценке ошибок систем искусственного интеллекта (ИИ), учитывающий их полную стоимость и специфические свойства алгоритмических ошибок. **Методология.** Исследование сочетает интегративный обзор литературы по экономике ИИ и машинному обучению с учётом издержек, анализ девяти задокументированных случаев ошибок алгоритмических систем в 2012–2024 гг. и экономико-математическое моделирование. **Результаты.** Предложена классификация издержек ошибок ИИ, включающая прямые, операционные, правовые, рыночные и социальные издержки. Выделены пять свойств, отличающих ошибки алгоритмов от ошибок человека: масштабируемость, коррелированность, непрозрачность, запаздывание обнаружения и самоусиление. Разработана модель ожидаемой стоимости ошибок и чистого эффекта внедрения ИИ с надбавкой за системный риск. Иллюстративный расчёт для кредитного скоринга показал, что алгоритм снижает потери почти на 40 %, однако незамеченный дрейф данных делает эффект внедрения отрицательным, а гибридная схема с участием человека обеспечивает наименьшие совокупные издержки. **Оригинальность.** Статистические показатели качества модели связаны с полной экономической стоимостью

ошибок, а коррелированность алгоритмических ошибок введена как источник системного риска. **Практическая значимость.** Результаты применимы при обосновании внедрения ИИ, настройке порогов решений, организации мониторинга моделей и разработке режимов ответственности.

Ключевые слова: искусственный интеллект, ошибки алгоритмов, стоимость ошибки, принятие решений с учётом издержек, дрейф данных, системный риск, ответственность за ИИ.

Abstract. Purpose. The study develops an approach to the economic assessment of errors made by artificial intelligence (AI) systems that captures their full cost and the specific properties of algorithmic errors. **Design/methodology/approach.** The paper combines an integrative review of research on the economics of AI and cost-sensitive machine learning, an analysis of nine documented cases of algorithmic failures in 2012–2024, and economic modelling. **Findings.** A classification of AI error costs is proposed, comprising direct, operational, legal, market, and social costs. Five properties distinguish algorithmic errors from human errors: scalability, correlation, opacity, delayed detection, and self-reinforcement. A model of the expected cost of errors and the net benefit of AI adoption with a systemic risk premium is developed. An illustrative credit scoring calculation shows that an algorithm reduces losses by almost 40%, yet undetected data drift turns the net effect of adoption negative, while a hybrid human-in-the-loop scheme yields the lowest total cost. **Originality/value.** The study links statistical performance metrics to the full economic cost of errors and introduces error correlation as a source of systemic risk. **Practical implications.** The results inform AI adoption decisions, decision threshold calibration, model monitoring, and the design of liability regimes.

Keywords: artificial intelligence, algorithmic errors, cost of error, cost-sensitive decision-making, data drift, systemic risk, AI liability.

Введение

Искусственный интеллект (ИИ) всё чаще принимает или готовит решения, от которых зависят деньги, здоровье и судьбы людей: одобрение кредитов, отбор кандидатов на работу, выявление мошенничества, оценка недвижимости, назначение социальных выплат, ответы клиентам от имени компании. Экономическая привлекательность таких систем обычно обосновывается ростом скорости и точности решений и снижением затрат. Значительно реже в расчётах учитывается обратная сторона — стоимость ошибок. Между тем любая система, основанная на статистическом прогнозировании, неизбежно ошибается: вопрос состоит не в том, будут ли ошибки, а в том, как часто они возникают, какие формы принимают и во что обходятся.

История последних лет содержит множество примеров того, насколько дорогими могут быть ошибки алгоритмов. Компания Zillow после сбоя алгоритмической модели оценки жилья была вынуждена списать сотни миллионов долларов и полностью закрыть направление бизнеса (Zillow Group, 2021). Автоматизированная система взыскания долгов по социальным выплатам в Австралии завершилась мировым соглашением на сумму около 1,2 млрд австралийских долларов и работой королевской комиссии (Royal Commission into the Robodebt Scheme, 2023). В Нидерландах скандал вокруг алгоритмического выявления мошенничества с пособиями на детей привёл к отставке правительства в январе 2021 года (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020). Ошибка чат-бота в демонстрационном ролике новой ИИ-системы Google совпала с падением рыночной капитализации Alphabet примерно на 100 млрд долларов за один день (Reuters, 2023).

Цель настоящей статьи — разработать подход к экономической оценке ошибок систем ИИ, учитывающий их прямые и косвенные издержки, а также

специфические свойства алгоритмических ошибок, отличающие их от ошибок человека. Исследование направлено на поиск ответов на следующие вопросы.

RQ1. Из каких компонентов складывается полная экономическая стоимость ошибок систем ИИ?

RQ2. Чем ошибки алгоритмов отличаются от ошибок человека с экономической точки зрения?

RQ3. Каким образом стоимость ошибок может быть учтена при принятии решения о внедрении ИИ и настройке его параметров?

Вклад работы состоит в следующем. Во-первых, предложена классификация издержек ошибок ИИ, включающая прямые, операционные, правовые, рыночные и социальные издержки. Во-вторых, на основе анализа девяти задокументированных случаев выделены специфические свойства алгоритмических ошибок: масштабируемость, коррелированность, непрозрачность, запаздывание обнаружения и самоусиление. В-третьих, разработана модель ожидаемой стоимости ошибок алгоритмической системы и чистого экономического эффекта от её внедрения с учётом риска коррелированных ошибок. В-четвёртых, на условном примере кредитного скоринга показано, что незамеченный дрейф данных способен полностью уничтожить экономический эффект от внедрения ИИ, а гибридная схема с участием человека обеспечивает наименьшие совокупные издержки.

2. Теоретические основания

2.1. ИИ как технология прогнозирования

А. Агравал, Дж. Ганс и А. Голдфарб предложили рассматривать современный ИИ прежде всего как технологию, радикально удешевляющую прогнозирование (Agrawal et al., 2018). В их подходе решение складывается из прогноза (оценки вероятностей возможных исходов) и суждения (оценки выгод и потерь от каждого исхода). ИИ удешевляет прогноз, но не заменяет суждения, и потому возрастает ценность комплементарного ему ресурса —

способности правильно оценивать последствия действий (Agrawal et al., 2019). Этот подход непосредственно связан с проблематикой настоящей статьи: стоимость ошибки определяется не только тем, насколько часто ошибается прогноз, но и тем, насколько корректно в систему заложена оценка последствий разных ошибок.

2.2. Типы ошибок и принятие решений с учётом издержек

В задачах классификации различают ошибки первого и второго рода. Ложноположительная ошибка возникает, когда алгоритм ошибочно относит объект к целевому классу: отказывает надёжному заёмщику, блокирует добросовестную транзакцию, отклоняет подходящего кандидата. Ложноотрицательная ошибка возникает, когда алгоритм не распознаёт целевой объект: одобряет кредит будущему неплательщику, пропускает мошенническую операцию, не выявляет заболевание. Принципиально важно, что экономическая стоимость этих ошибок, как правило, асимметрична.

Ч. Элкан обосновал подход к обучению с учётом издержек, в рамках которого оптимальное решение определяется не максимизацией точности, а минимизацией ожидаемых потерь с учётом матрицы стоимости ошибок (Elkan, 2001). Из этого подхода следует, что оптимальный порог принятия решения зависит от соотношения стоимостей ошибок разных типов. Модель, оптимизированная по стандартной метрике точности, в экономическом смысле может быть далека от оптимальной, если стоимость ошибок разных типов существенно различается.

3. Методология

Исследование основано на сочетании трёх методов.

Первый метод — интегративный обзор литературы в области экономики ИИ, машинного обучения с учётом издержек, поведенческих исследований взаимодействия человека и автоматики, а также управления рисками. Поиск осуществлялся в базах Scopus, Web of Science и Google Scholar по ключевым

словам «cost of AI errors», «cost-sensitive learning», «algorithmic failure», «automation bias», «AI liability», «model drift».

Второй метод — анализ множественных случаев. В выборку включены девять случаев ошибок алгоритмических систем за 2012–2024 годы, отвечающих трём критериям: наличие публично задокументированных последствий (финансовые потери, судебные или регуляторные решения, официальные расследования); разнообразие сфер применения (частный бизнес, финансовые рынки, государственное управление, здравоохранение, клиентский сервис, право); разнообразие типов ошибок. Источниками послужили официальная отчётность компаний, решения судов и регуляторов, отчёты парламентских и королевских комиссий, а также рецензируемые научные публикации. Для каждого случая определялись тип ошибки, её экономические последствия и механизм возникновения издержек.

Третий метод — экономико-математическое моделирование. На основе теории принятия решений с учётом издержек построена модель ожидаемой стоимости ошибок алгоритмической системы, а её свойства продемонстрированы на условном примере с анализом альтернативных сценариев.

Ограничением методологии является то, что выборка случаев смещена в сторону крупных публичных инцидентов, поскольку именно они получают документальное отражение. Мелкие, но многочисленные ошибки, совокупная стоимость которых может быть сопоставимой или даже большей, как правило, остаются незамеченными. Этот эффект, однако, лишь усиливает вывод о том, что стоимость ошибок ИИ систематически недооценивается.

4. Результаты

4.1. Структура издержек ошибок ИИ

Прямые издержки, которые обычно учитываются при оценке качества модели, представляют лишь видимую часть совокупной стоимости ошибок.

Правовые, рыночные и социальные издержки, как правило, проявляются с запаздыванием, а их масштаб может многократно превышать прямые потери. Особую группу образуют социальные издержки, которые ложатся не на компанию, применяющую алгоритм, а на третьи лица и общество: пострадавших граждан, дискриминируемые группы, клиентов, лишённых доступа к услугам. С экономической точки зрения это отрицательные внешние эффекты, которые без регулирования не учитываются в решениях о внедрении ИИ.

4.2. Анализ случаев

Анализ случаев позволяет выявить несколько механизмов возникновения издержек. Случай Zillow иллюстрирует эффект неблагоприятного отбора. Алгоритм оценки стоимости жилья формировал предложения о покупке домов, но продавцы, располагающие лучшей информацией о своих объектах, принимали прежде всего завышенные предложения. В условиях меняющегося рынка это привело к систематической переплате: компания сообщила о списании стоимости запасов жилья примерно на 304 млн долларов в третьем квартале 2021 года, объявила о закрытии направления и сокращении около четверти персонала (Zillow Group, 2021). Этот случай показывает, что ошибка модели в рыночной среде не является случайной: контрагенты целенаправленно используют её слабости, и ошибки концентрируются в наиболее дорогих для компании сделках.

Случаи Robodebt и нидерландского скандала с пособиями демонстрируют специфику ошибок в государственном управлении. В обоих случаях автоматизированные системы применялись к большому числу граждан, а ошибка носила не случайный, а системный характер, затрагивая целые категории людей. Издержки включали не только компенсации и судебные расходы, но и подрыв доверия к государственным институтам, а в нидерландском случае — политический кризис. Существенно, что бремя

доказывания фактически перекладывалось на граждан, которые должны были оспаривать решения алгоритма.

Случаи Google Bard, Air Canada и Mata v. Avianca связаны с генеративным ИИ и демонстрируют новые формы издержек. Ошибка в рекламном ролике повлекла масштабную рыночную переоценку компании. Решение канадского трибунала по делу Air Canada, обязавшее авиакомпанию выплатить пассажиру около 812 канадских долларов за неверную информацию, предоставленную чат-ботом, невелико по сумме, но создало важный прецедент: компания несёт ответственность за сведения, сообщаемые её ИИ-ассистентом, и не может ссылаться на то, что чат-бот является «отдельным субъектом» (Moffatt v. Air Canada, 2024). В деле Mata v. Avianca суд наложил санкции на адвокатов, представивших процессуальные документы с вымышленными судебными прецедентами, сгенерированными ChatGPT (Mata v. Avianca, Inc., 2023).

Наконец, случаи Amazon и алгоритма в здравоохранении США иллюстрируют издержки смещённых данных. Экспериментальная система отбора резюме Amazon, обученная на исторических данных о найме, дискриминировала кандидатов-женщин, и компания отказалась от её использования (Dastin, 2018). Медицинский алгоритм систематически недооценивал потребности чернокожих пациентов (Obermeyer et al., 2019). В обоих случаях алгоритм не «ошибался» в статистическом смысле — он точно воспроизводил закономерности обучающих данных, — однако эти закономерности отражали историческую несправедливость.

5. Модель экономической оценки ошибок алгоритмической системы

Отвечая на RQ3, предлагается модель, связывающая статистические характеристики алгоритма с экономической стоимостью его ошибок. Рассмотрим систему, принимающую за период N бинарных решений. Ожидаемая стоимость ошибок системы определяется выражением:

$$E[C] = N \cdot [\pi \cdot FNR \cdot C_{FN} + (1-\pi) \cdot FPR \cdot C_{FP}] + C_S \quad (1)$$

где E[C] — ожидаемые совокупные издержки ошибок за период; N — число решений; π — доля целевых объектов (например, будущих неплательщиков) в совокупности; FNR — доля ложноотрицательных ошибок среди целевых объектов; FPR — доля ложноположительных ошибок среди нецелевых объектов; C_{FN} и C_{FP} — полная стоимость одной ошибки соответствующего типа, включающая все группы издержек из таблицы 1; C_S — затраты на функционирование системы, включая мониторинг и проверку решений.

В соответствии с подходом Элкана (Elkan, 2001), если модель выдаёт вероятность p принадлежности объекта к целевому классу, оптимальный порог, минимизирующий ожидаемые издержки, равен:

$$p^* = \frac{C_{FP}}{C_{FP} + C_{FN}} \quad (2)$$

Объект следует относить к целевому классу, если p > p*. Из формулы (2) следует, что при высокой стоимости ложноотрицательной ошибки по сравнению с ложноположительной оптимальный порог смещается вниз, и система должна действовать более осторожно, даже ценой большего числа ложных срабатываний.

Практическое значение формулы (2) нередко недооценивается. Многие системы классификации по умолчанию используют порог 0,5, то есть относят объект к целевому классу, если вероятность его принадлежности к этому классу превышает вероятность противоположного исхода. Такой порог оптимален лишь в частном случае равенства стоимостей ошибок обоих типов. В большинстве экономических задач это условие не выполняется: пропуск мошеннической операции обходится дороже, чем проверка добросовестной, а одобрение кредита неплательщику — дороже, чем отказ надёжному заёмщику. Использование порога по умолчанию в таких задачах означает

систематическое принятие решений с завышенными ожидаемыми потерями даже при высокой статистической точности модели. Кроме того, стоимость ошибок не остаётся постоянной: она меняется вместе с процентными ставками, ценами, регуляторными требованиями и репутационной чувствительностью клиентов. Следовательно, порог решения должен рассматриваться как управляемый экономический параметр, подлежащий периодическому пересмотру, а не как технический параметр модели, заданный разработчиком один раз.

Чистый экономический эффект от замены решений человека решениями алгоритма с учётом риска коррелированных ошибок предлагается оценивать как:

$$\Delta W = E[C_H] - E[C_{AI}] - C_I - \lambda \sigma_{AI} \quad (3)$$

где ΔW — чистый экономический эффект внедрения; E[C_H] и E[C_{AI}] — ожидаемые издержки ошибок при решениях человека и алгоритма соответственно; C_I — затраты на внедрение и сопровождение системы, приведённые к периоду; σ_AI — показатель дополнительного разброса потерь, обусловленного коррелированностью ошибок алгоритма; λ — коэффициент неприятия риска организации. Слагаемое λσ_AI отражает «надбавку за системный риск»: даже при равенстве ожидаемых потерь организация, не склонная к риску, должна предпочитать решения с меньшей вероятностью катастрофических исходов.

Для иллюстрации модели рассмотрим условный пример кредитной организации, рассматривающей 100 тыс. заявок в год. Доля будущих неплательщиков составляет 10 %, ожидаемые потери от одобрения кредита неплательщику — 300 тыс. руб., упущенная прибыль от отказа надёжному заёмщику — 40 тыс. руб. Оптимальный порог по формуле (2) составляет около 0,118: заявку следует отклонять, если прогнозная вероятность невозврата превышает 11,8 %. Сопоставим четыре сценария: решения принимают

кредитные специалисты; решения принимает алгоритм; алгоритм работает в условиях незамеченного дрейфа данных, снизившего его точность; применяется гибридная схема, при которой алгоритм принимает решения автоматически, а около 30 % неоднозначных заявок направляется на проверку специалисту стоимостью 500 руб. за заявку. Параметры точности заданы условно и выбраны для демонстрации свойств модели. Результаты расчёта представлены в таблице 3 и на рисунке 1.

Источник: расчёт автора по формулам (1) и (3). N = 100 тыс. заявок, π = 0,10, C_{FN} = 300 тыс. руб., C_{FP} = 40 тыс. руб.; ΔW рассчитан относительно решений специалистов при C_I = 150 млн руб. и λσ_AI = 0.

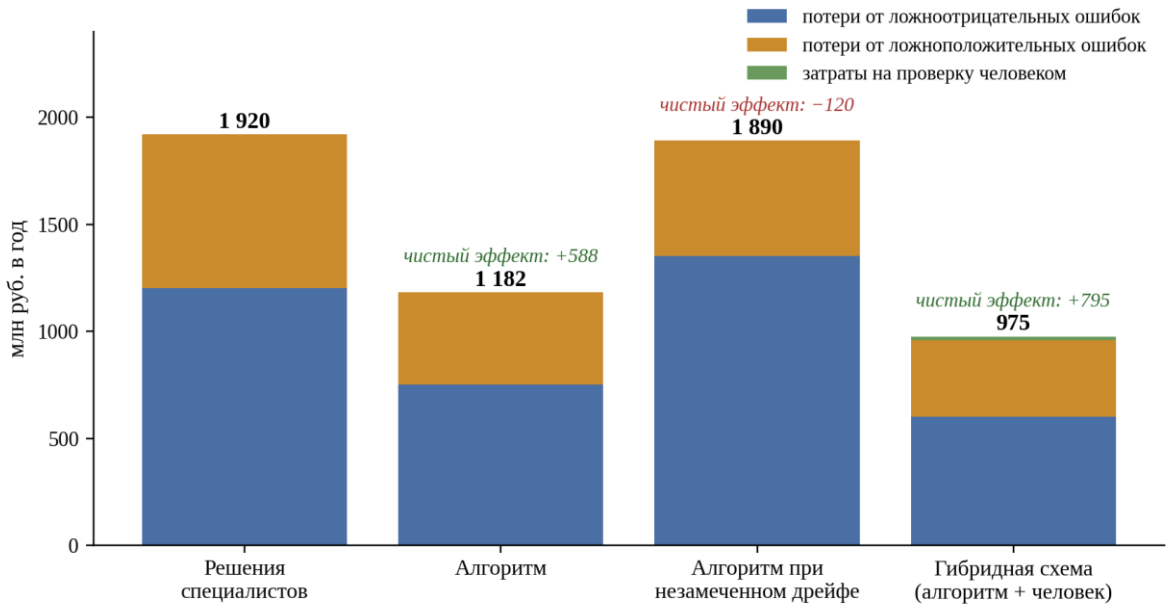


Рисунок 1. Структура ожидаемых издержек ошибок по сценариям (условный пример)

Источник: расчёт автора. Чистый эффект указан относительно решений специалистов с учётом затрат на систему 150 млн руб. в год.

Расчёт позволяет сделать несколько выводов. Во-первых, при корректной работе алгоритм снижает ожидаемые издержки ошибок с 1 920 до 1 182 млн руб. в год, то есть почти на 40 %. При затратах на внедрение и сопровождение системы, условно принятых равными 150 млн руб. в год,

чистый эффект составляет около 588 млн руб. Во-вторых, незамеченный дрейф данных, повысивший долю ложноотрицательных ошибок до 45 %, увеличивает издержки до 1 890 млн руб., что практически равно уровню решений человека. С учётом затрат на систему чистый эффект становится отрицательным и составляет около −120 млн руб.: внедрение ИИ без мониторинга качества оборачивается убытком. В-третьих, гибридная схема обеспечивает наименьшие совокупные издержки — около 975 млн руб., включая 15 млн руб. затрат на проверку специалистами, а её чистый эффект, около 795 млн руб., превышает эффект полной автоматизации. Следует подчеркнуть, что гибридная схема эффективна лишь при условии, что специалисты действительно проверяют решения, а не подтверждают их механически под влиянием автоматизационной предвзятости.

На основе модели и анализа случаев могут быть сформулированы предложения для последующей эмпирической проверки. P1: организации, настраивающие пороги решений с учётом асимметрии стоимости ошибок, достигают меньших совокупных потерь, чем организации, оптимизирующие стандартные метрики точности. P2: чем выше коррелированность ошибок алгоритма и чем больше задержка между решением и наблюдаемым результатом, тем выше экономическая ценность непрерывного мониторинга модели. P3: гибридные схемы принятия решений снижают совокупные издержки по сравнению с полной автоматизацией тем сильнее, чем выше асимметрия стоимости ошибок и неопределённость среды. P4: установление ответственности компаний за решения их ИИ-систем снижает долю социальных издержек в совокупной стоимости ошибок за счёт их интернализации.

6. Обсуждение

Исследование развивает экономический анализ ИИ в трёх направлениях. Во-первых, оно дополняет подход, рассматривающий ИИ как технологию

удешевления прогнозов (Agrawal et al., 2018), явным анализом издержек ошибочных прогнозов и показывает, что экономический эффект ИИ определяется не только точностью модели, но и корректностью заложенной в систему оценки последствий ошибок. Во-вторых, работа переводит проблематику ошибок из плоскости статистических метрик в плоскость экономических потерь, связывая подход Элкана (Elkan, 2001) с полной классификацией издержек, включая правовые, рыночные и социальные. В-третьих, выделение свойства коррелированности алгоритмических ошибок позволяет применить к оценке внедрения ИИ логику управления портфельным риском: алгоритм может снижать ожидаемые потери, одновременно увеличивая хвостовые риски.

Полученные результаты уточняют и вывод Клейнберга и соавторов о превосходстве алгоритмов над решениями человека (Kleinberg et al., 2018). Это превосходство реально, но условно: оно сохраняется лишь при стабильности среды, наличии мониторинга и корректной настройке порогов. При нарушении этих условий преимущество может исчезнуть или смениться убытком, как показал сценарий дрейфа данных.

Отдельного внимания заслуживает формирование рынка услуг по управлению ошибками ИИ. К нему относятся независимый аудит алгоритмов, сервисы мониторинга качества моделей, а также страхование ответственности за ошибки ИИ, которое начинают предлагать крупные страховщики. С экономической точки зрения развитие этого рынка выполняет две функции. Во-первых, оно позволяет распределить риски коррелированных ошибок, которые для отдельной компании могут оказаться катастрофическими. Во-вторых, страховщики и аудиторы формируют рыночную оценку стоимости ошибок, устанавливая премии и требования к качеству систем в зависимости от их рискованности, и тем самым создают стимулы к инвестициям в

надёжность, аналогичные тем, которые страхование создаёт в других отраслях.

Для регуляторов ключевой вывод состоит в необходимости интернализации социальных издержек ошибок ИИ. Европейский союз движется в этом направлении: Регламент об искусственном интеллекте устанавливает требования к системам высокого риска (European Parliament and Council, 2024a), а новая Директива об ответственности за дефектную продукцию прямо распространяет режим ответственности на программное обеспечение, включая системы ИИ (European Parliament and Council, 2024b). Решение по делу Air Canada демонстрирует, что и в отсутствие специального законодательства суды склонны возлагать ответственность за ошибки ИИ на компанию, которая его применяет (Moffatt v. Air Canada, 2024).

7. Ограничения и направления будущих исследований

Исследование имеет ряд ограничений. Во-первых, анализ случаев основан на выборке крупных публичных инцидентов, что не позволяет оценить распределение и совокупную стоимость многочисленных мелких ошибок. Во-вторых, предложенная модель предполагает бинарный характер решений и известную стоимость ошибок; в реальных задачах решения могут быть многовариантными, а стоимость ошибок — неопределённой и зависящей от контекста. В-третьих, параметры иллюстративного расчёта заданы условно, и количественные выводы требуют проверки на реальных данных. В-четвёртых, методика оценки надбавки за системный риск требует дальнейшей разработки.

Перспективными направлениями будущих исследований являются эмпирическая проверка предложений P1–P4 на данных финансовых организаций, разработка методов оценки коррелированности ошибок алгоритмов и её влияния на хвостовые риски, исследование эффективности различных схем участия человека в алгоритмических решениях с учётом

автоматизационной предвзятости, а также анализ влияния режимов ответственности на инвестиции компаний в качество и безопасность систем ИИ.

8. Заключение

Ошибки систем искусственного интеллекта неизбежны, и их экономическая стоимость является неотъемлемой частью экономики ИИ. Проведённое исследование показало, что полная стоимость ошибок складывается из прямых, операционных, правовых, рыночных и социальных издержек, причём видимые прямые потери нередко составляют лишь незначительную её часть. Анализ задокументированных случаев — от Zillow и Knight Capital до Robodebt, Air Canada и алгоритмов в здравоохранении — продемонстрировал, что ошибки алгоритмов обладают специфическими свойствами: масштабируемостью, коррелированностью, непрозрачностью, запаздыванием обнаружения и самоусилением.

Предложенная модель связывает статистические характеристики алгоритма с экономической стоимостью его ошибок и позволяет оценить чистый эффект внедрения ИИ с учётом надбавки за системный риск. Иллюстративный расчёт показал, что корректно настроенный алгоритм способен существенно снизить потери по сравнению с решениями человека, однако незамеченный дрейф данных может полностью уничтожить этот эффект, а гибридная схема с участием человека обеспечивает наименьшие совокупные издержки.

Главный вывод исследования состоит в том, что экономическая ценность ИИ определяется не безошибочностью алгоритма, а способностью организации управлять его ошибками: правильно оценивать их стоимость, настраивать решения с учётом асимметрии потерь, своевременно выявлять сбои и сохранять человеческий контроль там, где цена ошибки наиболее

высока. Вопрос «сколько стоит неправильное решение алгоритма» должен задаваться не после инцидента, а до внедрения системы.

Список литературы

Parlementaire ondervragingscommissie Kinderopvangtoeslag. (2020). *Ongekend onrecht* [Unprecedented injustice]. Tweede Kamer der Staten-Generaal.

Reuters. (2023, February 8). Alphabet shares dive after Google AI chatbot Bard flubs answer in ad. *Reuters*.

Royal Commission into the Robodebt Scheme. (2023). *Report*. Commonwealth of Australia.

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in Neural Information Processing Systems 28* (pp. 2503–2511). Curran Associates.

U.S. Securities and Exchange Commission. (2013). *In the matter of Knight Capital Americas LLC: Order instituting administrative and cease-and-desist proceedings* (Release No. 34-70694).

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ, 369*, m1328. https://doi.org/10.1136/bmj.m1328

Zillow Group. (2021, November 2). *Q3 2021 shareholder letter*. Zillow Group, Inc.

Альянс в сфере искусственного интеллекта. (2021). *Кодекс этики в сфере искусственного интеллекта*. https://ethics.a-ai.ru

Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (ред. от 15.02.2024). (2019). Доступ из справ.-правовой системы «КонсультантПлюс».