

**ТЕХНИКАЛЫҚ ҒЫЛЫМДАР / ТЕХНИЧЕСКИЕ НАУКИ /
ENGINEERING SCIENCES****QAZAQ JOURNAL OF YOUNG SCIENTIST****2026, Vol.4, No. 4 S (April)**<https://qazaqjournal.kz/>

УДК 004.89:658.5

**ИНТЕЛЛЕКТУАЛЬНАЯ МЕТОДИКА ОЦЕНКИ ЭФФЕКТИВНОСТИ
ВНЕДРЕНИЯ AI-РЕШЕНИЙ НА ОСНОВЕ МЕТРИКИ ITIS И МЕТОДОВ
МАШИННОГО ОБУЧЕНИЯ****Байтукенов Б. А.**

*магистрант, образовательная программа 7M06137 «IT-Менеджмент»,
Казахский университет технологии и бизнеса им. К. Кулажанова, г. Астана,
Республика Казахстан*

Научный руководитель: Шайханова А.К., PhD, ассоциированный
профессор, Казахский университет технологии и бизнеса им. К. Кулажанова, г.
Астана, Республика Казахстан

В статье представлена авторская методика оценки эффективности внедрения решений искусственного интеллекта для компаний малого и среднего бизнеса, основанная на интегральной метрике AI Impact Score (ITIS) и методах машинного обучения. Разработанная модель БАК («Бизнес-цель — AI-решение — KPI») формализует связь между функциональными модулями AI-платформы и бизнес-результатами компаний-клиентов. Метрика ITIS рассчитывается как среднегеометрическое четырёх параметров: бизнес-эффекта, стратегического веса, покрытия бизнес-целей и зрелости реализации. Для прогнозирования класса эффекта внедрения проведён сравнительный эксперимент четырёх алгоритмов классификации на датасете UCI Bank Marketing (45 211 наблюдений); в качестве базовой модели выбран Random Forest Classifier (Accuracy = 0,844; Macro F1 = 0,730). Апробация методики на данных казахстанского стартапа AI Pradavan дала итоговый ITIS = 0,90, соответствующий классу «высокоэффективное внедрение».

Ключевые слова: искусственный интеллект, оценка эффективности ИТ, AI Impact Score, ITIS, машинное обучение, случайный лес, цифровая трансформация, SaaS-стартап.

Введение

Цифровая трансформация бизнеса перестала быть стратегической альтернативой и превратилась в одно из ключевых условий конкурентоспособности компаний. Вместе с тем вопрос об измеримой ценности ИТ-инвестиций остаётся методологически открытым. По данным McKinsey, значительная часть организаций уже внедряет решения на основе искусственного интеллекта, однако получение устойчивого и измеримого эффекта от таких инициатив по-прежнему сопряжено с существенными трудностями [1]. Этот разрыв, который можно интерпретировать как проявление «IT Value Gap», особенно заметен в сегменте AI-решений, где традиционные подходы к оценке эффективности ИТ, включая финансовые и управленческие модели, обнаруживают ограничения при работе с нематериальными и многокомпонентными результатами [4], [7], [10]–[12].

Применительно к сегменту малого и среднего бизнеса и SaaS-стартапов данная проблема усугубляется рядом обстоятельств. Во-первых, финансовые методы оценки требуют монетизации нефинансовых эффектов, таких как скорость ответа, качество клиентского взаимодействия, уровень автоматизации и зрелость данных, что далеко не всегда корректно с методологической точки зрения. Во-вторых, сбалансированная система показателей (BSC) ориентирована преимущественно на корпоративный контур управления и длительный горизонт внедрения, что ограничивает её применимость к быстроразвивающимся цифровым продуктам [7]. В-третьих, существующие модели оценки бизнес-ценности ИТ и информационных систем не в полной мере учитывают модульную архитектуру современных AI-платформ, в которых разные компоненты по-разному влияют на итоговый бизнес-результат [4], [11], [12]. Кроме того, для практик внедрения искусственного интеллекта характерна необходимость связывать операционные эффекты с прикладными сценариями использования, что подчёркивается и в исследованиях по реализации AI в реальном бизнес-контексте [13], [14].

Цель настоящего исследования заключается в разработке и апробации интегральной методики оценки эффективности внедрения AI-решений, объединяющей авторскую метрику AI Impact Score (ITIS) с методами машинного обучения для ретроспективной и прогнозной оценки. Объектом апробации выступает казахстанская B2B SaaS-платформа AI Pradavan — решение класса Conversational AI для автоматизации входящих продаж в мессенджерах с интеграциями amoCRM и Bitrix24 [15].

Научная новизна работы состоит в следующем: во-первых, предложена интегральная метрика оценки эффективности AI-платформ, адаптированная к специфике SaaS-стартапов и МСБ; во-вторых, разработана трёхуровневая модель БАК, формализующая связь между бизнес-целью клиента, функциональным модулем AI-решения и измеримым KPI; в-третьих, реализован алгоритм расчёта ITIS на основе среднегеометрической агрегации, исключающей компенсацию критически низких значений; в-четвёртых, проведена экспериментальная верификация выбора базовой модели машинного обучения на независимом наборе данных [5], [8].

Материалы и методы исследования

Методологической основой исследования выступает синтез нескольких теоретических подходов к оценке ценности ИТ и выравниванию технологий с бизнес-целями. В качестве первой концептуальной опоры используется Strategic Alignment Model Хендерсона и Венкатрамана, согласно которой максимальная ценность ИТ достигается при согласовании технологической и бизнес-стратегии [2]. Вторая опора представлена методологией OKR, обеспечивающей декомпозицию стратегических целей в измеримые результаты [3]. Третьим основанием служит концепция IT Business Value, разграничивающая транзакционные, информационные и трансформационные эффекты использования ИТ [4]. Дополнительный контекст для проектирования оценочной модели обеспечивают подходы к измерению успешности информационных систем и их организационного эффекта [11], [12].

Разработанная модель БАК реализует трёхуровневую иерархическую конструкцию. Уровень 1 предполагает формализацию бизнес-целей компании-клиента в SMART-формате. Уровень 2 описывает матрицу соответствия функциональных модулей AI-платформы этим целям. Уровень 3 включает KPI-индикаторы, характеризующие достижение каждой цели с привязкой к конкретному модулю. Такая архитектура обеспечивает атрибуцию результата: для каждого наблюдаемого улучшения KPI прослеживается цепочка «KPI → модуль → бизнес-цель», что отличает предложенный подход от более агрегированных управленческих моделей, включая BSC [7], [11].

Интегральная метрика AI Impact Score (ITIS) для каждого функционального модуля i рассчитывается как среднегеометрическое четырёх нормализованных параметров:

$$ITIS_i = (E_i \times W_i \times C_i \times M_i)^{(1/4)}, \quad (1)$$

где E_i — нормализованный бизнес-эффект модуля; W_i — стратегический вес; C_i — коэффициент покрытия бизнес-целей; M_i — показатель зрелости реализации. Все параметры принадлежат интервалу $[0; 1]$. Итоговый ITIS платформы вычисляется как средневзвешенное по модулям:

$$ITIS_{\text{итог}} = \Sigma(ITIS_i \times W_i) / \Sigma W_i. \quad (2)$$

Принципиальное методологическое решение — выбор среднегеометрической агрегации вместо среднеарифметической. Среднегеометрическое обладает свойством «компенсационного барьера»: критически низкое значение одного параметра не может быть замаскировано высокими значениями остальных. Например, при $E = 0,95$, $W = 0,95$, $C = 0,95$ и $M = 0,10$ среднеарифметическое даст 0,74, создавая ложное впечатление зрелости, тогда как среднегеометрическое даст 0,52, корректно отражая ограниченность реализации.

Для прогнозирования класса эффекта внедрения задача формулируется как многоклассовая классификация $f: X_i \rightarrow \text{Class}(ITIS_i)$, где X_i — вектор характеристик компании-клиента (отрасль, размер компании, объём входящих заявок, среднее время ответа, наличие CRM, число каналов коммуникации, уровень цифровой зрелости и др.); $\text{Class}(ITIS_i) \in \{0, 1, 2\}$ — класс ожидаемого эффекта (низкий, средний, высокий) по шкале: $ITIS < 0,45$ — низкий; $0,45 \leq ITIS < 0,70$ — средний; $ITIS \geq 0,70$ — высокий.

Сравнительный эксперимент четырёх алгоритмов классификации (логистическая регрессия, дерево решений, случайный лес, градиентный бустинг) проведён на открытом датасете UCI Bank Marketing [5], содержащем 45 211 наблюдений и 16 признаков с целевой переменной — отклик клиентов на маркетинговую кампанию. Методика эксперимента: стратифицированное разбиение на обучающую и тестовую выборки в соотношении 80/20; предобработка через Pipeline (SimpleImputer + StandardScaler для числовых признаков, SimpleImputer + OneHotEncoder для категориальных); стратифицированная 5-кратная перекрёстная проверка. Параметры моделей: Logistic Regression — class_weight="balanced", max_iter=2000; Decision Tree — max_depth=6, class_weight="balanced"; Random Forest — n_estimators=200, max_depth=10, class_weight="balanced"; Gradient Boosting — n_estimators=150, learning_rate=0,1, max_depth=3. Все эксперименты выполнены с random_state=42. Ключевая метрика оценки — Macro F1, устойчивая к дисбалансу классов[8].

Апробация методики проведена на данных AI Pradavan за 2024–2025 гг. Выборка — агрегированные данные 50+ клиентов из семи отраслей (ритейл, NoReCa, строительство, образование, туризм, автомобильный бизнес, услуги). Применён метод «до/после» (before-after comparison) с экспертной оценкой весовых коэффициентов командой разработчика. Для обеспечения достоверности проведена тройная верификация результата через независимые источники данных: публичные видеозвоны клиентов, динамика роста клиентской базы, отраслевые бенчмарки Glassix [6], [15].

Результаты и обсуждение

Результаты сравнительной оценки четырёх алгоритмов машинного обучения представлены в таблице 1. Как видно из таблицы 1 и рисунков 1–3, наилучшие результаты по совокупности метрик продемонстрировали ансамблевые методы. На рисунке 1 отражены показатели моделей по метрикам перекрёстной проверки, на рисунке 2 — результаты на тестовой выборке, а рисунок 3 позволяет наглядно сопоставить распределение показателей качества по всем рассматриваемым моделям.

Таблица 1 — Сравнительная оценка алгоритмов машинного обучения

Алгоритм	Accuracy (CV)	Macro F1 (CV)	Accuracy (Test)	Macro F1 (Test)
Logistic Regression	0,840	0,722	0,839	0,718
Decision Tree	0,811	0,692	0,796	0,680
Random Forest	0,847	0,732	0,844	0,730
Gradient Boosting	0,906	0,731	0,906	0,729

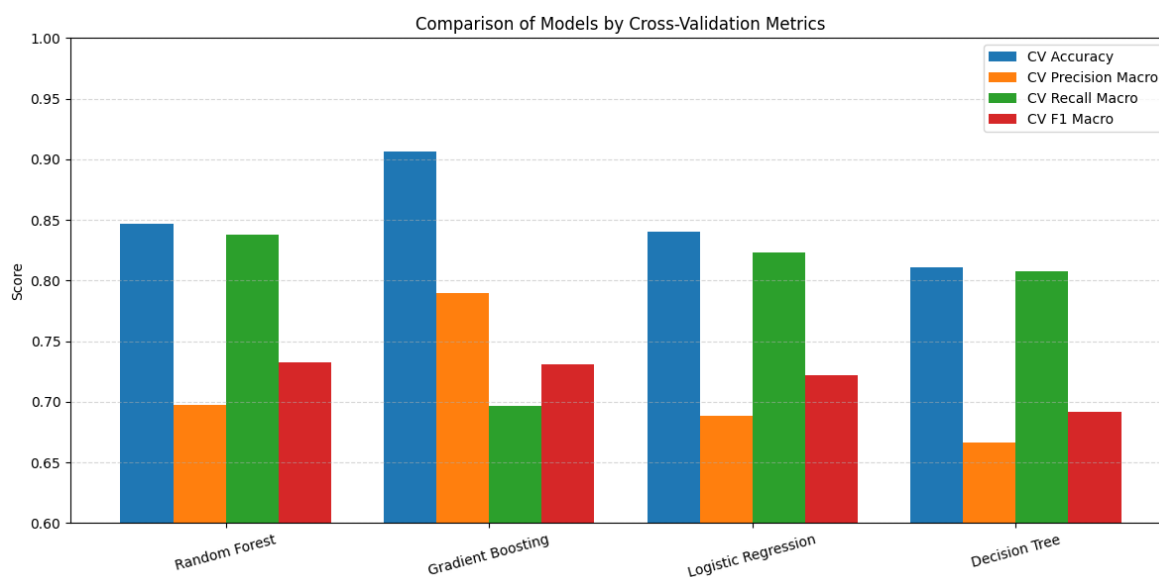


Рисунок 1 — Сравнение моделей машинного обучения по метрикам перекрёстной проверки

Анализ полученных результатов свидетельствует о преимуществе ансамблевых методов над линейными моделями и одиночным деревом решений. Логистическая регрессия демонстрирует удовлетворительное качество классификации, однако уступает ансамблям по значению Macro F1, что может указывать на нелинейный характер зависимости между характеристиками клиента и классом эффекта внедрения. Дерево решений показало наименьшую устойчивость: разница между значениями CV Macro F1

и Test Macro F1 составляет 0,012, что свидетельствует о склонности модели к переобучению.

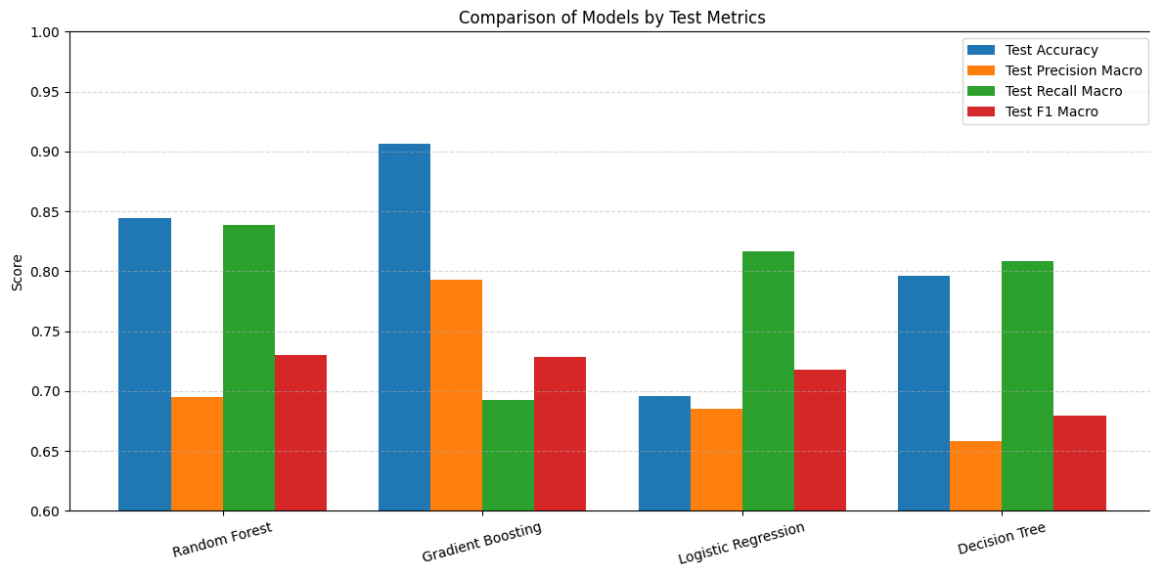


Рисунок 2 — Сравнение моделей машинного обучения по метрикам на тестовой выборке

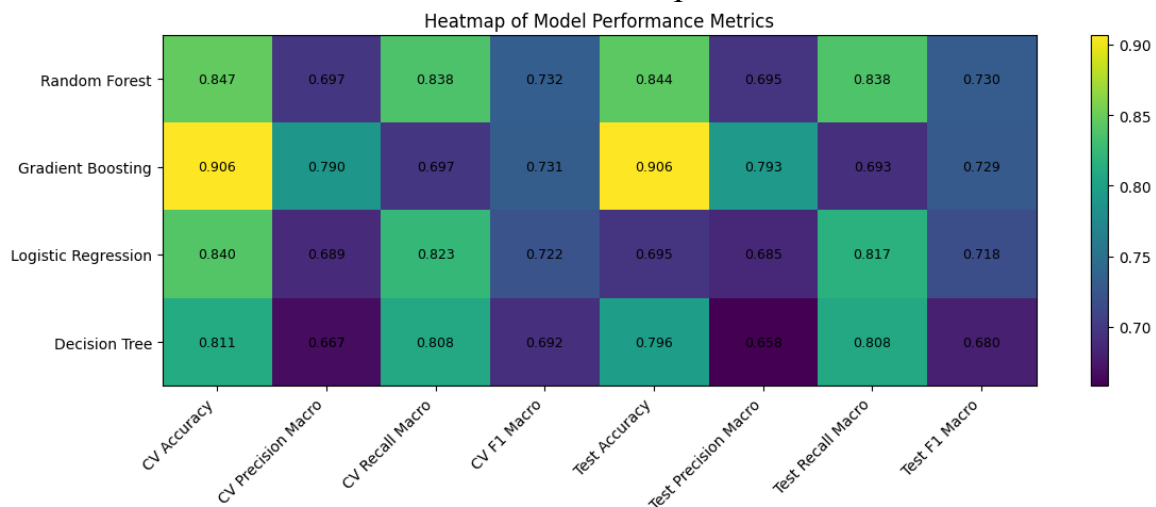


Рисунок 3 — Тепловая карта показателей качества моделей машинного обучения

Наилучшие результаты по совокупности метрик продемонстрировали Random Forest и Gradient Boosting. Gradient Boosting обеспечивает более высокое значение Accuracy (0,906 против 0,844), однако Random Forest незначительно превосходит его по Macro F1 (0,730 против 0,729), которая является ключевой метрикой в условиях дисбаланса классов. С учётом совокупности критериев, включая значение Macro F1, устойчивость результатов, интерпретируемость и применимость при ограниченном объёме данных, в качестве базовой модели был выбран Random Forest Classifier.

Дополнительным преимуществом данной модели является возможность оценки важности признаков (feature importance), что повышает объяснимость результатов и практическую применимость модели.

Результаты апробации методики на платформе AI Pradavan представлены в таблице 2.

Таблица 2 — Расчёт AI Impact Score по модулям платформы AI Pradavan

Модуль AI Pradavan	E	W	C	M	ITIS _i
Модуль мгновенного NLP-ответа	0,95	0,30	0,90	0,85	0,92
CRM-интеграция (amoCRM, Bitrix24)	0,88	0,25	0,85	0,90	0,87
Оmnikanальная платформа	0,90	0,20	0,80	0,88	0,88
No-code конструктор ботов	0,92	0,15	0,95	0,80	0,91
Аналитика и отчётность	0,82	0,10	0,75	0,78	0,83
Итоговый взвешенный ITIS	—	1,00	—	—	0,90

Примечание — рассчитано авторами на основе данных AI Pradavan за 2024–2025 гг.

Как следует из данных таблицы 2, итоговое значение взвешенного показателя ITIS составило 0,90, что позволяет отнести платформу AI Pradavan к классу высокоэффективных AI-решений. Полученный результат свидетельствует о том, что внедрение платформы обеспечивает не только выраженные операционные эффекты, но и более глубокие изменения в организации процессов продаж компаний-клиентов. К числу наиболее значимых эффектов относятся сокращение времени первого ответа на 99 %, снижение нагрузки на менеджеров на 80 % и переход от реактивной модели обработки обращений к проактивной автоматизированной модели с непрерывной квалификацией лидов в режиме 24/7 [6], [15].

Анализ значений ITIS по отдельным функциональным модулям показывает, что наименьшее значение показателя зафиксировано для модуля «Аналитика и отчётность» (ITIS = 0,83). Это обусловлено сравнительно более низкими значениями коэффициента покрытия бизнес-целей (C = 0,75) и показателя зрелости реализации (M = 0,78). Полученные результаты позволяют сделать вывод, что именно данный модуль выступает основным ограничивающим элементом в структуре совокупной эффективности платформы и, следовательно, представляет собой приоритетное направление дальнейшего продуктового развития.

Дополнительным подтверждением валидности полученного результата выступает его верификация по трём независимым источникам. Во-первых,

публичные отзывы клиентов указывают на наличие устойчивых изменений в модели продаж после внедрения платформы [15]. Во-вторых, рост клиентской базы с примерно 10 до более чем 50 компаний в течение 12 месяцев при сохранении подписочной модели косвенно свидетельствует о достижении Product-Market Fit и практической состоятельности продукта, что согласуется с работами по внедрению искусственного интеллекта в прикладных бизнес-процессах [13], [14]. В-третьих, зафиксированный прирост конверсии у клиентов AI Pradavan на уровне 45 % превышает медианное значение, представленное в отраслевых исследованиях Glassix (23 %) [6], что дополнительно подтверждает высокий уровень практической эффективности рассматриваемого решения.

Выводы

По результатам проведённого исследования установлено, что разработанная методика оценки эффективности внедрения AI-решений позволяет системно учитывать ключевые методологические ограничения традиционных подходов к оценке ИТ-инвестиций. Проблема атрибуции решается за счёт трёхуровневой архитектуры модели БАК, обеспечивающей связь каждого изменения КРІ с конкретным функциональным модулем AI-платформы. Проблема агрегации преодолевается посредством использования среднегеометрической формулы ITIS, обладающей свойством «компенсационного барьера» и предотвращающей некорректное усреднение разнородных эффектов. Проблема временного сдвига учитывается через включение параметра зрелости реализации M, позволяющего разграничивать устойчивые эффекты и результаты ранней стадии внедрения.

Сравнительный эксперимент по выбору базовой модели машинного обучения показал, что Random Forest Classifier является наиболее обоснованным решением с точки зрения совокупности критериев качества, устойчивости и интерпретируемости. При значениях Accuracy = 0,844 и Macro F1 = 0,730 данная модель продемонстрировала наилучший баланс между точностью классификации и способностью корректно работать в условиях дисбаланса классов, что обеспечивает воспроизводимость и прикладную надёжность предложенных методологических решений.

Апробация методики на данных платформы AI Pradavan показала её практическую применимость. Полученное итоговое значение ITIS = 0,90 позволяет отнести рассматриваемую платформу к классу высокоэффективных AI-решений. Дополнительная верификация результата по нескольким независимым источникам подтверждает достоверность сделанных выводов и свидетельствует о том, что предложенная методика может использоваться как инструмент диагностики эффективности внедрения AI-решений в прикладных условиях.

Практическая значимость исследования заключается в возможности применения разработанной методики в нескольких направлениях. Для

разработчиков AI-платформ она может выступать инструментом выявления функциональных ограничений и обоснования приоритетов продуктового развития. Для компаний-клиентов методика представляет интерес как средство обоснования инвестиционных решений в области цифровой трансформации. На отраслевом уровне предложенный подход может служить основой для формирования сопоставимых бенчмарков эффективности внедрения AI-решений, что приобретает особую актуальность в контексте реализации Концепции развития искусственного интеллекта Республики Казахстан на 2024–2029 годы.

Вместе с тем исследование имеет ряд ограничений. К ним относятся ограниченный объём эмпирической выборки, включающей данные более чем 50 клиентов, а также сравнительно небольшой горизонт наблюдения, охватывающий период 2024–2025 гг., что не позволяет в полной мере оценить долгосрочные трансформационные эффекты внедрения. Перспективы дальнейших исследований связаны с автоматизацией калибровки весовых коэффициентов на основе метода анализа иерархий (АНР), расширением ML-компонента за счёт применения методов кластеризации, а также анализа тональности клиентских диалогов с использованием моделей класса BERT. Дополнительным направлением является внешняя валидация предложенной методики на расширенной выборке, включающей более 200 практических кейсов.

Список литературы

1. McKinsey & Company. The state of AI in early 2024: Gen AI adoption spikes and starts to generate value [Электронный ресурс]. 2024. URL: McKinsey (дата обращения: 15.03.2025).
2. Henderson J. C., Venkatraman N. Strategic alignment: Leveraging information technology for transforming organizations // IBM Systems Journal. 1993. Vol. 32, No. 1. P. 472–484. DOI: 10.1147/sj.382.0472.
3. Doerr J. Measure What Matters: How Google, Bono, and the Gates Foundation Rock the World with OKRs. New York: Portfolio, 2018. 320 p.
4. Kohli R., Grover V. Business value of IT: An essay on expanding research directions to keep up with the times // Journal of the Association for Information Systems. 2008. Vol. 9, No. 1. P. 23–39. DOI: 10.17705/1jais.00147.
5. Moro S., Cortez P., Rita P. A data-driven approach to predict the success of bank telemarketing // Decision Support Systems. 2014. Vol. 62. P. 22–31. DOI: 10.1016/j.dss.2014.03.001.
6. Glassix. Study shows AI chatbots enhance conversion by 23% [Электронный ресурс]. URL: Glassix (дата обращения: 20.03.2025).

7. Kaplan R. S., Norton D. P. The Balanced Scorecard: Translating Strategy into Action. Boston: Harvard Business School Press, 1996. 322 p.
8. Breiman L. Random forests // Machine Learning. 2001. Vol. 45, No. 1. P. 5–32. DOI: 10.1023/A:1010933404324.
9. Об утверждении Концепции развития искусственного интеллекта на 2024–2029 годы: Постановление Правительства Республики Казахстан от 24 июля 2024 года № 592 [Электронный ресурс]. URL: Эділет (дата обращения: 15.03.2025).
10. Brynjolfsson E., Hitt L. M. Beyond computation: Information technology, organizational transformation and business performance // Journal of Economic Perspectives. 2000. Vol. 14, No. 4. P. 23–48. DOI: 10.1257/jep.14.4.23.
11. DeLone W. H., McLean E. R. The DeLone and McLean model of information systems success: A ten-year update // Journal of Management Information Systems. 2003. Vol. 19, No. 4. P. 9–30. DOI: 10.1080/07421222.2003.11045748.
12. Melville N., Kraemer K., Gurbaxani V. Review: Information technology and organizational performance: An integrative model of IT business value // MIS Quarterly. 2004. Vol. 28, No. 2. P. 283–322.
13. Davenport T. H., Ronanki R. Artificial intelligence for the real world // Harvard Business Review. 2018. Vol. 96, No. 1. P. 108–116.
14. Ransbotham S., Kiron D., Gerbert P., Reeves M. Reshaping business with artificial intelligence: Closing the gap between ambition and action. MIT Sloan Management Review; Boston Consulting Group, 2017.
15. AI Pradavan [Электронный ресурс]: официальный сайт. URL: <https://aipradavan.city-innovation.kz> (дата обращения: 20.03.2026).

INTELLIGENT METHODOLOGY FOR ASSESSING THE EFFECTIVENESS OF AI SOLUTION IMPLEMENTATION BASED ON THE ITIS METRIC AND MACHINE LEARNING METHODS

Baitukenov B. A.

Scientific Supervisor: A.K. Shaikhanova

The article presents the author's methodology for assessing the effectiveness of artificial intelligence solution implementation for small and medium-sized businesses, based on the integral AI Impact Score (ITIS) metric and machine learning methods. The developed BAK model ("Business goal — AI solution — KPI") formalises the relationship between the functional modules of an AI platform and the business results of client companies. The ITIS metric is calculated as the geometric mean of four parameters: business effect, strategic weight, business goal coverage, and

implementation maturity. To forecast the class of implementation effect, a comparative experiment of four classification algorithms was conducted on the UCI Bank Marketing dataset (45,211 observations); the Random Forest Classifier was chosen as the baseline model (Accuracy = 0.844; Macro F1 = 0.730). The methodology was tested on the data of the Kazakhstani AI Pradavan startup and yielded a final ITIS = 0.90, corresponding to the "highly effective implementation" class.

Keywords: artificial intelligence, IT effectiveness assessment, AI Impact Score, ITIS, machine learning, random forest, digital transformation, SaaS startup.

ITIS МЕТРИКАСЫ МЕН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНЕ НЕГІЗДЕЛГЕН АІ-ШЕШІМДЕРДІ ЕНГІЗУ ТИІМДІЛІГІН БАҒАЛАУДЫҢ ИНТЕЛЛЕКТУАЛДЫ ӘДІСТЕМЕСІ

Байтукенов Б. А.

Ғылыми жетекші: Шайханова А.К.

Мақалада шағын және орта бизнес компаниялары үшін жасанды интеллект шешімдерін енгізу тиімділігін бағалаудың авторлық әдістемесі ұсынылған. Әдістеме AI Impact Score (ITIS) интегралдық метрикасына және машиналық оқыту әдістеріне негізделген. Әзірленген БАК моделі («Бизнес-мақсат — АІ-шешім — КРІ») АІ платформасының функционалдық модульдері мен клиент-компаниялардың бизнес нәтижелері арасындағы байланысты формалдайды. ITIS метрикасы төрт параметрдің геометриялық орташа мәні ретінде есептеледі: бизнес-әсер, стратегиялық салмақ, бизнес-мақсаттарды қамту деңгейі және іске асырудың жетілу дәрежесі. Енгізу әсерінің класын болжау үшін UCI Bank Marketing деректер жиынтығында (45 211 бақылау) төрт классификация алгоритміне салыстырмалы эксперимент жүргізілді. Негізгі модель ретінде Random Forest Classifier таңдалды (Accuracy = 0,844; Macro F1 = 0,730). Әдістемені қазақстандық AI Pradavan стартапының деректері негізінде апробациялау нәтижесінде ITIS көрсеткіші 0,90-ға тең болып, «жоғары тиімді енгізу» класына сәйкес келетіні анықталды.

Кілт сөздер: жасанды интеллект, IT тиімділігін бағалау, AI Impact Score, ITIS, машиналық оқыту, кездейсоқ орман, цифрлық трансформация, SaaS-стартап