

ОБЪЯСНИМЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В КАДРОВОЙ АНАЛИТИКЕ: АНАЛИЗ ФАКТОРОВ ПРОИЗВОДИТЕЛЬНОСТИ

А. Н. КОЗИНЕЦ¹⁾

¹⁾Белорусский государственный университет информатики и радиоэлектроники,
ул. П. Бровки, 6, 220013, г. Минск, Беларусь

Аннотация. В условиях цифровизации и роста значения человеческого капитала в кадровой аналитике особую актуальность приобретают модели объяснимого искусственного интеллекта (*explainable artificial intelligence*). С использованием градиентной модели *XGBoost* анализируется влияние надпрофессиональных навыков и цифровых компетенций на ключевые показатели эффективности сотрудников. Для интерпретации вклада признаков используется метод *SHAP*, обеспечивающий прозрачность прогнозов. Разрабатывается методология, применяемая на этапах отбора открытых данных, генерации признаков и интерпретации модели. Устанавливается, что такие факторы, как коммуникабельность, цифровая грамотность и стаж, оказывают значимое влияние на результативность труда. Обсуждаются управленческие последствия применения объяснимого искусственного интеллекта в поддержке решений по развитию персонала. Подчеркивается важность интерпретируемых моделей для повышения доверия к алгоритмам в кадровой аналитике.

Ключевые слова: кадровая аналитика; объяснимый искусственный интеллект; надпрофессиональные навыки; цифровые компетенции; градиентный бустинг; метод *SHAP*; производительность сотрудников; ключевые показатели эффективности.

EXPLAINABLE ARTIFICIAL INTELLIGENCE IN HUMAN RESOURCE ANALYTICS: ANALYSIS OF PERFORMANCE FACTORS

A. N. KAZINETZ^a

^aBelarusian State University of Informatics and Radioelectronics,
6 P. Browki Street, Minsk 220013, Belarus

Abstract. In the context of digitalisation and the growing importance of human capital, explainable artificial intelligence models in human resource analytics are particularly relevant. Using the *XGBoost* gradient model, the impact of transversal skills and digital competencies on key employee performance indicators is analysed. The *SHAP* method, which ensures forecast transparency, is used to interpret the contribution of features. A methodology is developed for the stages of open data selection, feature generation, and model interpretation. It is established that factors such as communication skills,

Образец цитирования:

Козинец АН. Объяснимый искусственный интеллект в кадровой аналитике: анализ факторов производительности. Журнал Белорусского государственного университета. Экономика. 2025;2:30–41.
EDN: EJPJVH

For citation:

Kazinetz AN. Explainable artificial intelligence in human resource analytics: analysis of performance factors. *Journal of the Belarusian State University. Economics*. 2025;2:30–41. Russian. EDN: EJPJVH

Автор:

Александр Николаевич Козинец – аспирант кафедры экономики инженерно-экономического факультета. Научный руководитель – кандидат экономических наук, доцент В. А. Пархименко.

Author:

Aliaksandr N. Kazinets, postgraduate student at the department of economics, faculty of engineering and economics.
kazinets.science@gmail.com
<https://orcid.org/0009-0003-0083-3504>

digital literacy, and tenure have a significant impact on performance. The managerial consequences of using explainable artificial intelligence to support human resource development decisions are discussed. The importance of interpretable models for increasing trust in algorithms in human resource analytics is emphasised.

Keywords: human resource analytics; explainable artificial intelligence; transversal skills; digital competencies; gradient boosting; *SHAP* method; employee performance; key performance indicators; KPIs.

Введение

Технологии анализа данных уже радикально изменили кадровые процессы – от подбора персонала до его развития. Кадровая аналитика перестала быть вспомогательным сервисом и превратилась в стратегический инструмент, напрямую влияющий на операционную эффективность и конкурентоспособность компаний. В условиях нарастающей цифровизации трудовых процессов формируется колоссальный массив данных о работниках, который охватывает демографические параметры, траектории профессионального развития, результаты их профессиональной деятельности и поведенческие характеристики. Эффективное использование таких данных требует не только точного прогнозирования, но и ясного объяснения причин, на которых базируются аналитические выводы.

В этом контексте особую значимость приобретает парадигма *explainable artificial intelligence* (ХАИ) – объяснимого искусственного интеллекта (ИИ), представляющего собой совокупность методологических подходов, направленных на повышение интерпретируемости и прозрачности предсказательных моделей. Применение ХАИ критически важно в сфере управления человеческими ресурсами, где алгоритмически сформированные рекомендации напрямую влияют на карьерные траектории сотрудников и стратегические решения организаций. Этические, правовые и институциональные ограничения становятся неотъемлемыми элементами для разработки и внедрения систем ИИ в кадровый менеджмент. Согласно актуальному нормативному регулированию ЕС инструменты ИИ, применяемые в трудовой сфере, квалифицируются как технологии с повышенным уровнем регуляторной значимости, что обуславливает необходимость соблюдения принципов прозрачности, обязательной оценки воздействия и предотвращения потенциальной дискриминации¹. В 2024 г. вступил в силу Регламент 2024/1689 Европейского парламента и совета от 13 июня 2024 г. (*AI Act*) – первый общеевропейский нормативный акт, системно регулирующий использование ИИ, в соответствии с которым кадровые системы были признаны социально значимыми. Документ предусматривает обязательное ведение отчета об ответственности за использование ИИ и присвоение регистрационного идентификатора моделям, принимающим решения, которые способны повлиять на профессиональный путь человека. В государствах – участниках СНГ, несмотря на отсутствие специализированных актов, действуют универсальные правовые механизмы, обеспечивающие защиту персональных данных и недопущение дискриминации при обработке информации о работниках². В октябре 2021 г. Комиссия по равным возможностям при трудоустройстве США инициировала реализацию программы *Artificial Intelligence and Algorithmic Fairness*, направленной на снижение рисков системной необъективности при использовании инструментов ИИ в принятии кадровых решений³.

Настоящее исследование посвящено интеграции ХАИ в аналитические практики управления персоналом в целях выявления детерминант производственной результативности. Эмпирическая база сформирована на основе сведений из базы данных O*NET⁴ и набора данных WNS HR Analytics⁵, содержащего структурированную информацию о персонале крупной корпорации. В качестве прогностического инструмента использована градиентная модель *XGBoost*, подвергнутая гиперпараметрической оптимизации с помощью библиотеки *Optuna*⁶. Для декомпозиции вклада отдельных факторов в итоговое предсказание применен метод *SHAP*, обеспечивающий интерпретируемость на уровне индивидуальных решений модели.

¹Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 March 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // European Union : website. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689> (date of access: 03.05.2025).

²О защите персональных данных : Закон Респ. Беларусь, 7 мая 2021 г., № 99-3 // Нац. правовой интернет-портал Респ. Беларусь. URL: <https://pravo.by/document/?guid=12551&p0=H12100099> (дата обращения: 03.05.2025).

³EEOC Launches Initiative on artificial intelligence and algorithmic fairness // U. S. Equal Employment Opportunity Commission : website. URL: <https://www.eeoc.gov/newsroom/eeoc-launches-initiative-artificial-intelligence-and-algorithmic-fairness> (date of access: 03.05.2025).

⁴O*NET Resource Center : website. URL: <https://www.onetcenter.org/> (date of access: 03.05.2025).

⁵WNS HR analytics [Electronic resource] // Kaggle. URL: <https://www.kaggle.com/akbaldawa/hr-analytics> (date of access: 03.05.2025).

⁶Optimize your optimization: an open source hyperparameter optimization framework to automate hyperparameter search [Electronic resource] // Optuna. URL: <https://optuna.org/> (date of access: 03.05.2025).

Предложенный подход позволяет не только достичь высокой прогностической точности, но и оценить экономическую значимость факторов человеческого капитала, включая профессиональные навыки, накопленный опыт, образовательные траектории и поведенческие индикаторы. Модель рассматривается в рамках экономической теории производительности: проводится анализ того, какие характеристики работников обеспечивают наибольшую отдачу в организационном контексте. Дополнительно исследуется вопрос алгоритмической справедливости: осуществляется оценка наличия систематических отклонений в предсказаниях модели в зависимости от пола сотрудников и апробируются процедуры постобработки, направленные на нивелирование потенциальных гендерных искажений.

Теоретические основы исследования

В экономической теории, начиная с фундаментальных положений Г. С. Беккера о человеческом капитале, неоднократно подчеркивалась ключевая роль индивидуальных характеристик работников (уровня образования, профессионального стажа, квалификаций) в формировании их производственной результативности [1]. Однако в условиях цифровой трансформации экономики и повсеместного внедрения технологий обработки больших данных спектр учитываемых факторов значительно расширился: наряду с классическими индикаторами в аналитические модели все чаще стали включаться количественно измеримые параметры, отражающие мотивацию, поведенческие паттерны, личностные характеристики и специфику профессиональных навыков.

Современная кадровая аналитика прошла путь от элементарной описательной статистики до сложных предиктивных систем, позволяющих, например, прогнозировать текучесть кадров, идентифицировать сотрудников с высоким потенциалом, а также моделировать индивидуальные траектории развития. Одним из ключевых источников стандартизированных данных о профессиях и требованиях к навыкам является база данных O*NET, содержащая структурированную информацию о знаниях, умениях, когнитивных и физических способностях, а также характеристиках трудового контекста для широкого круга профессиональных наименований. Внедрение сведений из базы данных O*NET в модели управления человеческими ресурсами способствует более точной оценке относительной значимости различных компонентов человеческого капитала, а также аналитическому выявлению взаимосвязей между характеристиками работников и показателями их производственной результативности. Согласно ряду эмпирических исследований наблюдается устойчивый рост значимости цифровых и коммуникативных компетенций, особенно в условиях активной цифровой трансформации организационно-производственной среды [2]. Например, аналитика специалистов Брукингского института показала, что в 2002–2016 гг. доля рабочих мест с высокой интенсивностью цифровых задач выросла с 5 до 23 %; при этом рабочие места, требующие более высокого уровня цифровых навыков, в среднем оплачиваются существенно выше, чем должности с низкими требованиями к цифровой подготовке⁷. Актуальные аналитические отчеты из области цифровых компетенций фиксируют двукратный рост спроса на вакансии, предполагающие наличие цифровых навыков, в период с 2012 по 2022 г., а также прогнозируют потенциальные экономические потери Великобритании в размере 27,6 млрд фунтов стерлингов в случае сохранения текущего дефицита соответствующих специалистов⁸.

Распространение технологий машинного обучения в управлении персоналом привело к формированию нового исследовательского направления – аналитики персонала (*people analytics*), в рамках которого разрабатываются и внедряются модели прогнозирования увольнений, повышения по службе, оценки результативности сотрудников и других кадровых процессов. К популярным моделям относятся случайные леса, градиентный бустинг, глубокие нейронные сети. Однако использование высокоэффективных, но слабо интерпретируемых моделей порождает фундаментальную проблему доверия – так называемый эффект черного ящика [3], препятствующий широкому принятию машинных решений на основе методов машинного обучения в организациях.

Указанный вызов способствовал активному развитию подходов к ХАИ. Систематический метаобзор, охватывающий свыше 300 научных публикаций, позволяет классифицировать используемые подходы и подчеркивает, что метод *SHAP* входит в тройку наиболее эффективных методов по совокупному критерию согласованности и точности (*consistency + accuracy*) [4]. Интегративная ревизия более поздних исследований в области аналитики персонала дает возможность констатировать экспоненциальный рост числа публикаций, в которых метод *SHAP* применяется в сочетании с моделью *XGBoost* для решения задач по оценке и прогнозированию производительности [5]. Современная научная литература описывает

⁷Muro M., Maxim R., Whiton J. Digitalization and the American workforce // Brookings Institution : website. URL: https://www.brookings.edu/wp-content/uploads/2017/11/mpp_2017nov15_digitalization_full_report.pdf (date of access: 03.05.2025).

⁸Ma H., Kollydas K., Lyons M., Green A. Digital skills demand and shortages and their impact in UK regional economies // University of Birmingham : website. URL: <https://www.birmingham.ac.uk/research/centres-institutes/city-region-economic-development-institute/publications/2024/digital-skills-demand-and-shortages-and-their-impact-on-the-uk-regions> (date of access: 03.05.2025).

широкий спектр методов ХАИ – от изначально интерпретируемых моделей (например, деревьев решений и наборов правил) до постобъяснительных (*post-hoc*) подходов, предназначенных для интерпретации сложных алгоритмов. К последним относятся, в частности, методы *LIME*, *SHAP*, *Partial dependence plots*, а также новейшие постобъяснительные техники *Contrastive-XAI* и *CXPlain*, ориентированные на *what-if*– объяснение [6]. Особое внимание уделяется методу *SHAP*, разработанному С. М. Лундбергом и С.-И. Ли, который адаптирует концепцию кооперативных игр к задаче распределения вклада признаков в итоговое предсказание [7].

В формализованном виде значение *SHAP* для признака i определяется как его средний маржинальный вклад во множество всех возможных подмножеств признаков, не содержащих i , по формуле

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (f(S \cup \{i\}) - f(S)),$$

где n – общее число признаков; S – подмножество признаков, не содержащих i ; $f(S)$ – функция предсказания модели. Метод *SHAP* обеспечивает как глобальные оценки важности (средние значения *SHAP* по множеству наблюдений), так и локальные объяснения на уровне отдельных предсказаний, удовлетворяя при этом критериям теоретической обоснованности, таким как линейность, симметрия и монотонность.

Роль ХАИ в кадровой аналитике выходит далеко за рамки повышения интерпретируемости алгоритмов: он непосредственно влияет на качество управленческих решений, принимаемых в условиях неопределенности и высокой ответственности. Ошибки алгоритмической классификации – как ложноположительные (продвижение неподходящего кандидата), так и ложноотрицательные (игнорирование перспективного работника) – сопряжены с прямыми и косвенными экономическими издержками (от репутационных потерь и снижения продуктивности до роста текучести и возникновения внутренних дисбалансов в командах).

Прозрачные и интерпретируемые модели позволяют менеджерам по управлению персоналом принимать обоснованные и этически выверенные решения, а также инициировать адресные инвестиции в развитие наиболее значимых факторов успеха (например, в обучение тем навыкам, которые признаны моделью ключевыми). Более того, ХАИ предоставляет инструментарий для диагностики алгоритмических искажений (*bias*). При отсутствии механизмов контроля модель может использовать предикторы, связанные с демографическими характеристиками (например, полом и возрастом), что способно привести к воспроизводству исторически сложившихся структурных диспропорций и форм социальной дискриминации. Ряд исследований в области экономики труда подтверждают, что даже высокоточные алгоритмы склонны наследовать институциональные предвзятости [8]. Результаты недавних экспериментов свидетельствуют о том, что сотрудники считают решения, принимаемые непрозрачными алгоритмическими системами, на 35 % менее справедливыми по сравнению с решениями, основанными на прозрачных процедурах [9]. Согласно официальному стенографическому отчету Комиссии по обеспечению равных возможностей при трудоустройстве США работодателям настоятельно рекомендуется проводить проверку показателя соотношения неблагоприятного воздействия (*bias impact ratio*) для всех применяемых инструментов ИИ, в противном случае возможно применение по отношению к ним мер административного взыскания, включая штрафные санкции⁹. В этой связи особенно актуально совмещение моделей машинного обучения с механизмами коррекции и постобработки, обеспечивающими соблюдение нормативных требований и принципов справедливости (например, через применение квот, порогов равных возможностей или через корректировку предсказаний на основе справедливых критериев).

Настоящее исследование опирается на вышеописанные теоретические и эмпирические разработки, оно является концептуальным мостом между алгоритмическими методами обработки данных и экономическим анализом трудовых ресурсов. Используя одну из наиболее эффективных моделей градиентного бустинга для табличных данных *XGBoost* в сочетании с методом *SHAP*, мы обеспечиваем как высокую прогностическую точность, так и глубокую интерпретируемость результатов, что позволяет не только идентифицировать ключевые факторы продуктивности, но и оценить справедливость, обоснованность рекомендаций модели в контексте организационного управления.

Материалы и методы исследования

Эмпирическая база исследования сформирована путем интеграции двух взаимодополняющих открытых массивов данных. Первичным источником выступил корпоративный набор данных WNS HR Analytics, содержащий демографические характеристики работников (пол, возраст, уровень образования,

⁹Navigating employment discrimination in AI and automated systems: transcript of the public hearing // U. S. Equal Employment Opportunity Commission : website. URL: <https://www.eeoc.gov/newsroom/eeoc-launches-initiative-artificial-intelligence-and-algorithmic-fairness> (date of access: 03.05.2025).

макрорегион), организационные параметры (подразделение и канал найма), показатели профессионального опыта и результативности (стаж, число завершённых обучающих программ, средний балл обучения, оценка эффективности за предыдущий период, идентификатор достижения более 80 % ключевых показателей эффективности, количество наград и бинарный идентификатор факта карьерного продвижения). Дополнительная информация о значимости надпрофессиональных навыков была взята из справочника «Skills – Importance» базы данных O*NET; сопоставление осуществлялось по связке должность – функциональный блок. В результате каждому сотруднику присвоены нормированные индексы важности коммуникационных и цифровых навыков, что расширило признаковое пространство. Актуальность использования комбинированной базы подтверждается результатами прикладного эксперимента: включение переменных из базы данных O*NET в модель машинного обучения, ориентированную на последующее объяснение ее работы с помощью метода SHAP, повысило ее эффективность на 4 процентных пункта по метрике $F1$ ¹⁰.

Доступ к сопоставимым кадровым данным белорусских организаций ограничен действующим законом о защите персональных данных и внутренними регламентами компаний, поэтому использование публично воспроизводимого массива отвечает международным стандартам верификации кадровых моделей и обеспечивает независимую проверяемость результатов.

Целевая переменная $is_promoted$ представляет собой бинарный индикатор факта карьерного продвижения; положительный класс составляет 8,5 % наблюдений, что приводит к значительному дисбалансу выборки. Для минимизации влияния данной переменной применялись стратифицированная кросс-валидация и оптимизированные весовые коэффициенты классов.

Для прогнозирования использовалась модель *XGBoost* – градиентный бустинг над решающими деревьями, высокоэффективный для табличных структур данных [10]. Данный алгоритм способен моделировать сложные нелинейные взаимосвязи и учитывать взаимодействия признаков, одновременно предотвращая переобучение через встроенную регуляризацию. Задача формулируется как бинарная классификация: модель предсказывает вероятность принадлежности сотрудника с признаками X к классу высокой эффективности с помощью формулы

$$P(y = \text{High KPI} | X).$$

Внутренняя структура модели *XGBoost* предполагает построение ансамбля из K деревьев решений, итоговая оценка $\hat{p}_i$ получается через логистическое преобразование суммы выходов деревьев с помощью следующего уравнения:

$$\hat{p}_i = \sigma(z_i) = \frac{1}{1 + e^{-z_i}},$$

где $z_i = \sum_{k=1}^K T_k(X_i)$, $T_k(X_i)$ – предсказание k -го дерева для объекта X_i , а $\sigma(\cdot)$ – сигмоида. Обучение модели

осуществляется посредством минимизации логарифмической функции потерь с регуляризационными слагаемыми, которые штрафуют модель за избыточную сложность деревьев решений.

В целях повышения прогностической способности модели был осуществлен подбор гиперпараметров с использованием библиотеки Optuna, основанной на алгоритме Tree-structured Parzen estimator [11]. В процессе оптимизации рассматривались следующие параметры: глубина деревьев, скорость обучения (*learning rate*), коэффициенты регуляризации, количество деревьев, а также параметры подвыборки признаков и наблюдений. Целевой функцией являлась площадь под ROC-кривой, вычисляемая по результатам 5-кратной стратифицированной кросс-валидации на тренировочной части выборки (80 % данных использовались для обучения модели, 20 % данных – для тестирования). После около 50 итераций алгоритма подбора гиперпараметров в библиотеке Optuna была получена стабильная конфигурация с валидационным $AUC = 0,75$. На независимой тестовой выборке модель достигла аналогичных значений (0,74–0,75), что подтверждает ее устойчивость и обобщающую способность.

Главная метрика качества – ROC AUC (*area under the receiver operating characteristic curve*), отражающая вероятность того, что случайно выбранный представитель класса высокой эффективности получит более высокую модельную оценку, чем представитель класса низкой эффективности, вычислялась по формуле

$$AUC = \frac{1}{N_1 N_0} \sum_{\substack{i: y_i = 1 \\ j: y_j = 0}} \mathbb{I}[\hat{p}_i > \hat{p}_j],$$

где N_1, N_0 – количества положительных и отрицательных наблюдений соответственно, а $\mathbb{I}[\cdot]$ – индикатор выполнения условия. Помимо AUC, рассчитывались следующие бизнес-ориентированные метрики:

¹⁰Putka D. J., Dahlke J. A., Burke M. I., Rounds J., Lewis P. Using machine learning to develop occupational interest profiles and high-point codes for the O*NET system // O*NET Resource Center : website. URL: https://www.onetcenter.org/dl_files/ML_OIPs.pdf (date of access: 03.05.2025).

точность (*accuracy*), полнота (*recall*), точность положительных прогнозов (*precision*), сбалансированная точность (*balanced accuracy*). Также строилась калибровочная кривая, отражающая соответствие вероятностных предсказаний истинной частоте наступления события.

В условиях дисбаланса классов для этической обоснованности выводов был реализован этап пост-обработки. Во-первых, порог вероятности отнесения к классу высокой эффективности был установлен не на уровне 0,5, а с учетом максимизации F -меры и бизнес-ограничений по числу возможных повышений. Эмпирически определенный порог обеспечивал отбор 20–25 % лучших сотрудников. Во-вторых, проверена чувствительность модели к полу и обнаружен дисбаланс, который устранен путем установки независимых порогов классификации для каждой гендерной группы. Данный подход обеспечил паритет – равную долю положительных решений среди женщин и мужчин (около 45 % положительных решений в каждой группе) при незначительном снижении точности во имя соблюдения принципа недискриминации.

Для объяснения предсказаний модели использовался метод *SHAP*, реализованный с помощью модуля *TreeExplainer* из библиотеки *SHAP* на языке Python. Интерпретация проведена в двух аспектах:

1) в аспекте глобальной значимости признаков. На основе усредненных абсолютных значений *SHAP* определен список 20 наиболее значимых признаков. Сводный график *Beeswarm* отражает вклад каждого признака в предсказание модели в зависимости от его значения;

2) аспекте взаимодействия признаков. С помощью значений взаимодействия *SHAP* (*SHAP interaction values*) проанализированы эффекты взаимного влияния признаков, в частности среднего балла обучения и важности коммуникативных навыков. Построена диаграмма рассеяния, на которой по оси абсцисс отложены значения признака «средний балл обучения», по оси ординат – соответствующие значения *SHAP*, отражающие вклад признака в предсказание. Цветовая шкала отображает влияние признака, связанного с уровнем развития коммуникативных навыков. Такой график позволяет визуализировать вариативность вклада в зависимости от контекста профессиональных требований.

Весь анализ реализован на языке Python в среде *Google Colab*, результаты полностью воспроизводимы¹¹. В следующем разделе представлена интерпретация эмпирических находок сквозь призму их теоретической и прикладной значимости.

Результаты и их обсуждение

Оптимизированная модель *XGBoost* продемонстрировала устойчивую способность к классификации. На рис. 1 представлена ROC-кривая, построенная по результатам анализа данных тестовой выборки; площадь под кривой составила 0,75, что подтверждает высокое качество ранжирования.

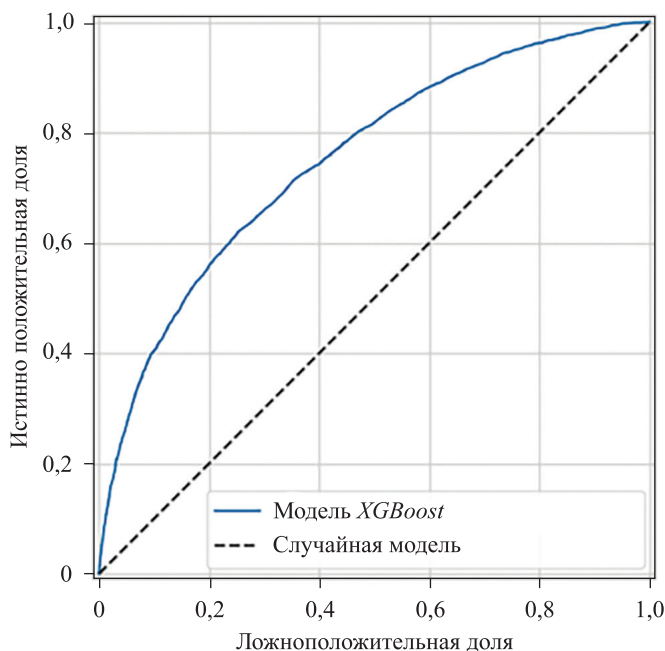


Рис. 1. ROC-кривая качества модели *XGBoost* (библиотека *Optuna*) на тестовой выборке (AUC = 0,75)

Fig. 1. ROC curve of the *XGBoost* model (*Optuna* library) quality on the test set (AUC = 0.75)

¹¹Explainable AI [Electronic resource] // Google Colaboratory. URL: <https://colab.research.google.com/drive/1bpuFE-0Gnw5F9-H0HPpT40ix6tm8OkqIY> (date of access: 03.05.2025).

Кривая существенно превосходит диагональ случайного угадывания, особенно в области низких уровней ложноположительных классификаций. При фиксированных 20 % ложноположительных срабатываний модель идентифицирует около 60 % сотрудников, относящихся к классу высокой эффективности. После постобработки выбранного порога классификации достигнуты следующие показатели: точность – около 70 %, полнота по классу высокой эффективности – около 65 %, полнота по классу низкой эффективности – около 71 %, что соответствует сбалансированной точности около 68 %.

На рис. 2 представлена нормализованная по строкам матрица ошибок, иллюстрирующая поведение модели после применения пороговой коррективы.

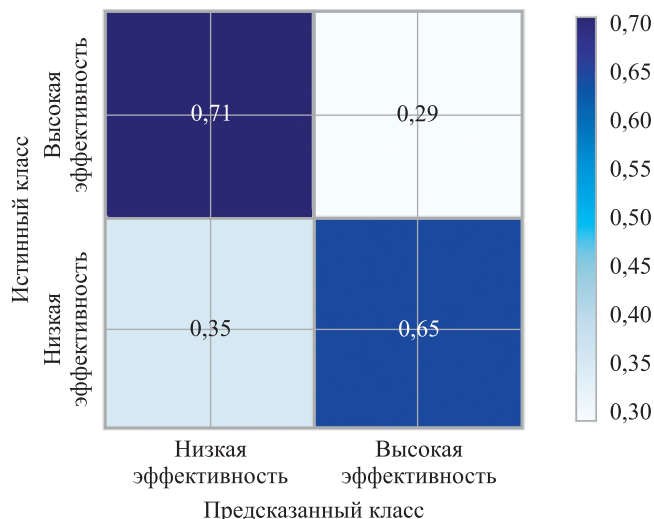


Рис. 2. Матрица ошибок модели *XGBoost* после постобработки порога

Fig. 2. Confusion matrix of the *XGBoost* model after threshold post-processing

Модель корректно классифицирует 71 % сотрудников, относящихся к классу низкой эффективности (*true negative*), и 65 % сотрудников, относящихся к классу высокой эффективности (*true positive*), ошибочно повышая показатели 29 % работников из первого класса и пропуская 35 % работников из второго класса. Доля невыявленных сотрудников, относящихся к классу высокой эффективности, объясняется целевой политикой компании: в условиях ограниченности карьерных возможностей модель приближенно воспроизводит актуальный механизм отбора. Калибровка вероятностей (рис. 3) показала адекватное соответствие предсказанных значений фактической частоте событий.

Например, для сотрудников с прогнозируемой вероятностью около 0,6 достижения высокой эффективности фактический процент успеха близок к 60 %. Такой результат подтверждает надежность вероятностных интерпретаций и позволяет менеджерам по управлению персоналом принимать решения с учетом количественных оценок неопределенности.

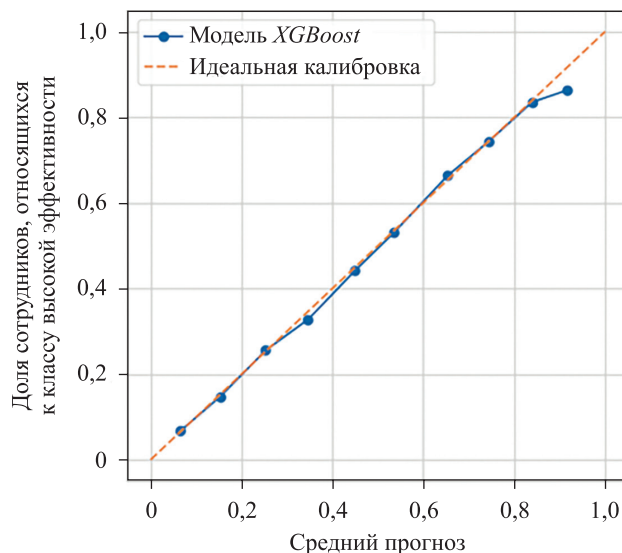


Рис. 3. Калибровочная кривая для модели *XGBoost*

Fig. 3. Calibration curve for the *XGBoost* model

Интерпретация значений *SHAP*. На рис. 4 представлен сводный график, который отражает 20 наиболее важных признаков по значениям, полученным в результате применения метода *SHAP* и интерпретируемым как вклад каждого признака в формирование предсказания модели.

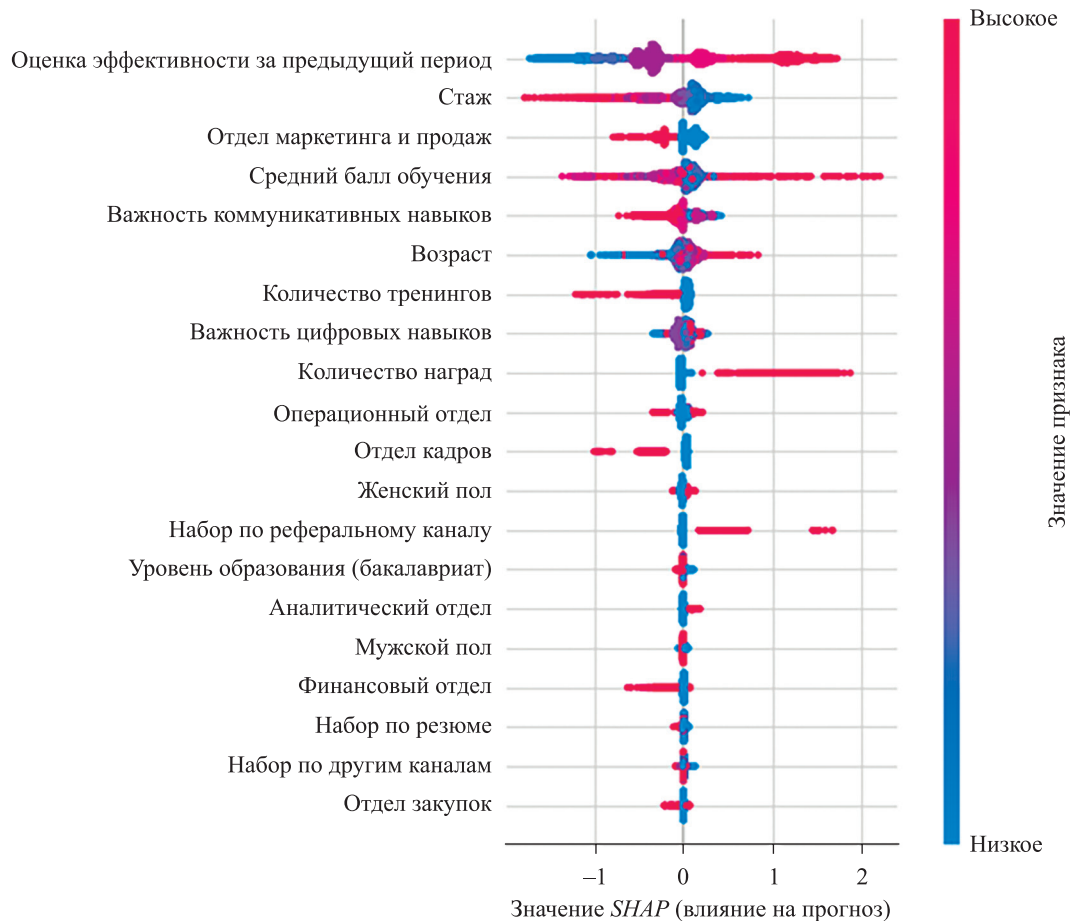


Рис. 4. Обзор 20 наиболее важных признаков, влияющих на прогноз высокой эффективности

Fig. 4. Top 20 feature importance overview for high performance forecasting

Наиболее информативным предиктором выступает оценка эффективности за предыдущий период, за которой следуют стаж, средний балл обучения, отдел, важность коммуникации, а также ряд демографических и поведенческих признаков. Вклад признаков интерпретируется по знаку и величине значений *SHAP*: положительные значения увеличивают вероятность отнесения сотрудников к классу высокой эффективности, отрицательные – снижают. В частности, стаж демонстрирует неоднозначное влияние, вероятно, вследствие нелинейной зависимости между продолжительностью работы и карьерным ростом.

Вклад функциональной принадлежности сотрудников (например, отдел маркетинга и продаж) также оказывается существенным: для отдела маркетинга и продаж, отдела кадров и операционного отдела отмечаются положительные смещения. Признаки, связанные с участием сотрудников в обучении и получении поощрений, подтверждают значимость нематериальной мотивации и профессионального развития. Присутствие признаков «мужской пол», «женский пол» и «уровень образования (бакалавриат)» в списке указывает на потенциальную чувствительность модели к социально значимым характеристикам, хотя влияние этих признаков умеренное.

Значения взаимодействия *SHAP*. Для детального анализа взаимодействий признаков, выявленных с помощью метода *SHAP*, построен рис. 5, иллюстрирующий модифицирующее влияние признака «важность коммуникативных навыков» на признак «средний балл обучения».

Виден значимый модифицирующий эффект: при высоком уровне коммуникационной значимости вклад результатов обучения в предсказание высокой эффективности сотрудника значительно увеличивается. Для коммуникационно-насыщенных ролей (например, работа с персоналом, продажи) высокие баллы обучения коррелируют с сильным положительным значением *SHAP*, тогда как при низкой важности коммуникативных навыков аналогичный результат оказывает слабое влияние. Таким образом, эффективность профессиональной деятельности зависит от развития надпрофессиональных навыков в конкретных функциональных условиях.

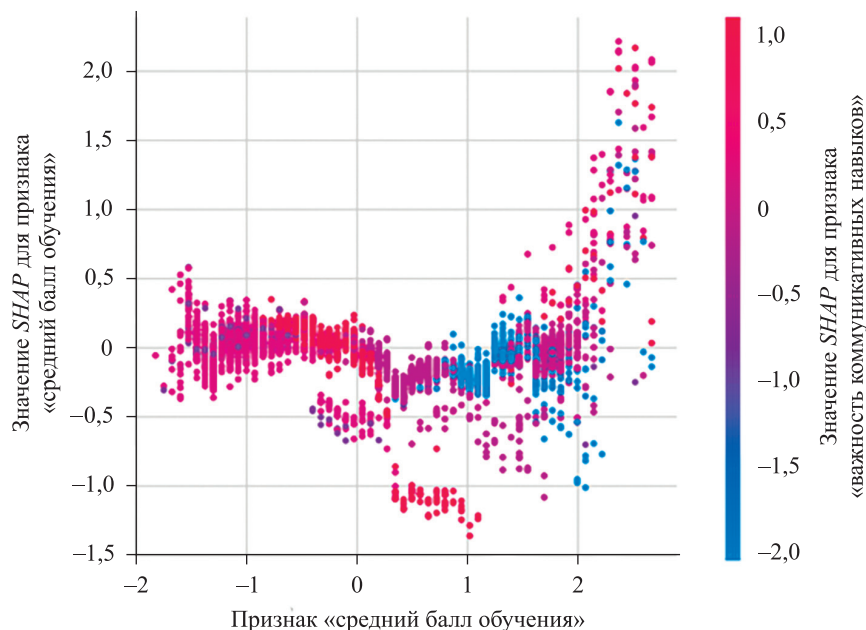


Рис. 5. Значение взаимодействия SHAP между признаками «средний балл обучения» и «важность коммуникативных навыков»

Fig. 5. SHAP interaction values between the «average training score» feature and «importance of communication» feature

Анализ гендерного баланса. На рис. 6 сопоставлены доли сотрудников, рекомендованных моделью к повышению, по признаку пола до и после постобработки.

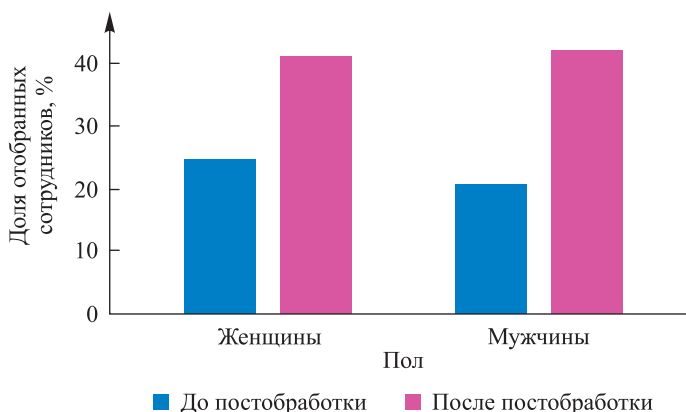


Рис. 6. Сравнение доли отобранных к повышению сотрудников по признаку пола

Fig. 6. Comparison of the proportion of employees selected for promotion by gender

До коррекции наблюдалась диспропорция: модель рекомендовала к продвижению около 25 % женщин и 20 % мужчин. После внедрения корректирующего механизма доли были выровнены до 45 % в каждой группе. Подобный подход соответствует принципу демографического паритета и представляет собой управленческое решение, направленное на обеспечение равных возможностей. Примененная коррекция привела к незначительному снижению точности положительных предсказаний, при этом повысив полноту для мужской группы без существенного влияния на интегральную метрику AUC и сбалансированную точность. Таким образом, инструментарий ХАИ позволил выявить латентные отклонения и осуществить адресную калибровку модели в соответствии с этическими приоритетами организации.

Результаты эмпирического анализа выявили ряд системно значимых детерминант производственной эффективности и подчеркнули критическую значимость ХАИ для принятия обоснованных управленческих решений в сфере управления персоналом. Таким образом, наши выводы соответствуют глобальному тренду цифровизации данной сферы: согласно опросу 4000 руководителей отделов по управлению персоналом доля компаний, применяющих инструменты ИИ в управлении персоналом, выросла с 58 до 72 % в 2024–2025 гг.¹²

¹²AI adoption among HR professionals rises to 72 % [Electronic resource] // Staffing Industry Analysts. URL: <https://www.staffing-industry.com/news/global-daily-news/ai-adoption-among-hr-professionals-rises-to-72> (date of access: 03.05.2025).

Опыт и кумулятивные эффекты. Предыдущая оценка эффективности оказалась ведущим прогностическим индикатором, что согласуется с теорией сигналов продуктивности и идеей инерционности трудовой результативности [12]. Высокая оценка за предыдущий период выступает сигналом наличия компетентности и мотивации, что экономически обоснованно в контексте карьерного продвижения. Аналогичным образом стаж работы демонстрирует преимущественно положительный вклад, однако вариативность значений *SHAP* указывает на эффект плато – возможное снижение продуктивности при длительной статичности в должности. От сферы управления персоналом это требует стратегических механизмов ротации и развития для сотрудников с большой выслугой лет.

Обучение и профессиональные компетенции. Высокий средний балл обучения, особенно в сочетании с высоким уровнем важности коммуникативных навыков, статистически существенно увеличивал вероятность отнесения сотрудника к классу высокой эффективности. Это обстоятельство эмпирически подтверждает инвестиционную гипотезу теории человеческого капитала: вложения в обучение трансформируются в прирост производительности [13]. Эффект взаимодействия подчеркивает необходимость дифференцированного подхода: для коммуникативно-ориентированных ролей (продажи, управление персоналом) эффективность обучения выражена ярче, чем для технических ролей. Следовательно, стратегическое распределение образовательных ресурсов по функциональным направлениям может повысить отдачу от корпоративного обучения.

Структурная мобильность и межфункциональные различия. Институциональная принадлежность демонстрирует значимое влияние на прогнозируемую эффективность. Признаки, отражающие принадлежность к отделу маркетинга и продаж, операционному отделу и отделу по управлению персоналом, ассоциированы с более высокой вероятностью карьерного роста. Эти различия, вероятно, обусловлены неравномерной структурой возможностей (числом должностных уровней, видимостью результатов, приоритетностью функций). Использование навыков, указанных в базе данных O*NET, позволило дополнительно категоризировать отделы по профилю компетенций (коммуникационные и цифровые отделы). Эта категоризация дает основу для разработки специализированных критериев оценки и планов развития персонала, соответствующих профилю задач.

Нематериальная мотивация. Наличие наград за достижения позитивно коррелирует с прогнозируемым уровнем эффективности, подтверждая значимость систем нематериального стимулирования. Такие награды можно рассматривать как ранние маркеры продуктивного поведения и использовать при формировании кадрового резерва.

Гендерная справедливость. Анализ показал, что женщины в выборке демонстрировали равную или несколько большую вероятность принадлежности к классу высокой эффективности, что противоречит гипотезе «стеклянного потолка» [14]. Это обстоятельство может быть результатом эффекта селекции: в организации остаются наиболее мотивированные и способные сотрудницы. Примененная корректировка порогов обеспечила паритет рекомендаций по полу (45 и 45 % сотрудников каждого пола), минимально повлияв на обобщающую способность модели. Данная мера соответствует критериям социальной справедливости и нормативным требованиям, в то время как методология ХАИ обеспечила прозрачность и контролируемость вмешательства. Параллельно информационное агентство *Reuters* указывает на растущие репутационные риски компаний, «спящих на ходу» в вопросах этичной автоматизации¹³.

Практическая применимость. Разработанная модель может быть интегрирована в корпоративные системы поддержки принятия решений, она позволяет формировать ранжированные списки сотрудников с объяснением причин рекомендаций. Этот факт повышает объективность и доверие к системе оценки, а также дает данные для обратной связи и оптимизации политик в сфере управления персоналом (от рекрутинга до обучения). Прогностическая аналитика становится инструментом стратегического управления человеческим капиталом.

Ограничения и перспективы. Исследование ограничено одним источником данных, отражающим структуру условной компании, что требует осторожности при обобщении. Контекстуальные различия между секторами (например, между технологическим и консалтинговым секторами) могут оказывать влияние на значимость разных факторов. Кроме того, анализ опирается на структурированные данные; включение неструктурированной информации (текстов, сетевых взаимодействий) расширит модель. Важно также рассматривать методы обеспечения справедливости как на этапе обучения модели (методы регуляризации, учитывающие справедливость), так и на этапе постобработки (оптимизационные подходы).

¹³Zappulla A. Comment: business leaders risk sleepwalking towards AI misuse [Electronic resource] // Reuters. URL: <https://www.reuters.com/sustainability/society-equity/comment-business-leaders-risk-sleepwalking-towards-ai-misuse-2024-11-19/> (date of access: 03.05.2025).

Полученные результаты демонстрируют синергетический эффект от объединения методов машинного обучения и экономической теории в контексте кадровой аналитики. Объяснимые модели обеспечивают одновременно точность прогнозов и этичность решений, что особенно важно в условиях цифровой трансформации трудовых отношений. Модель ХАИ становится не просто технологическим инструментом, а институтом, формирующим архитектуру справедливого и эффективного управления человеческим капиталом.

Заключение

Настоящее исследование продемонстрировало прикладную и методологическую состоятельность использования ХАИ в решении задач кадровой аналитики, направленных на выявление высокопродуктивных сотрудников и интерпретацию факторов, которые определяют уровень их результативности. Построенная модель *XGBoost*, оптимизированная посредством алгоритма *Optuna* и дополненная интерпретацией на основе метода *SHAP*, обеспечила не только высокую прогностическую точность ($AUC \approx 0,75$), но и теоретическую верифицируемость выводов.

Предыдущие достижения, накопленный стаж, успешность при обучении, наличие поощрений, функциональная принадлежность и требуемые компетенции – все это оказалось статистически значимыми детерминантами высоких эффективности. Эти результаты концептуально согласуются с экономическими теориями человеческого капитала и организационного поведения.

Использование метода *SHAP* позволило превратить модель из неинтерпретируемого алгоритмического механизма в прозрачный аналитический инструмент. Визуализация вклада признаков в индивидуальные прогнозы делает возможным рациональное принятие решений со стороны специалистов по управлению персоналом, снижая барьеры доверия к модели и увеличивая готовность к ее практическому внедрению.

Обнаруженное смещение по полу было устранено на этапе постобработки порогов, в результате чего обеспечен баланс рекомендаций (45 % мужчин и 45 % женщин). Такой результат демонстрирует, как ХАИ способствует соблюдению этических и нормативных стандартов, минимизируя риски алгоритмической дискриминации.

Объяснимость модели создала предпосылки для повышения точности кадровых решений, что снижает потери от управленческих ошибок (например, продвижение недостаточно квалифицированных сотрудников или игнорирование перспективных сотрудников). Кроме того, выявление значимых факторов позволяет целенаправленно инвестировать в развитие человеческого капитала и формировать кадровую политику на основе объективной обратной связи.

Объединяя точность алгоритмов и прозрачность интерпретации, ХАИ в кадровой аналитике формирует новую парадигму – управление персоналом на основе данных, подкрепленных теоретически осмысленными объяснениями. Проведенное исследование показало, что подобные модели не только применимы, но и необходимы для эффективного и справедливого принятия решений в условиях цифровой экономики. Перспективным направлением является распространение подходов, в рамках которых используется ХАИ, и на другие кадровые задачи: прогнозирование текучести, потребности в обучении; моделирование карьерных траекторий. При этически выверенном применении такие системы могут стать фундаментом для устойчивого повышения производительности как на уровне отдельных организаций, так и на уровне макроэкономической системы в целом.

Библиографические ссылки

1. Becker GS. *Human capital: a theoretical and empirical analysis*. Chicago: University of Chicago Press; 1964. 187 p.
2. Флек МБ, Угнич ЕА. Разработка модели цифровых компетенций работников в условиях цифровой трансформации предприятия. *Перспективы науки и образования*. 2023;3:706–723. DOI: 10.32744/pse.2023.3.43.
3. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. In: Balaji Krishnapuram, Mohak Shah, editors. *Proceedings of the 22nd ACM SIGKDD International conference on knowledge discovery and data mining; 13–17 August 2016; San Francisco, USA*. New York: Association for Computing Machinery; 2016. p. 1135–1144.
4. Arrieta AB, Diaz-Rodriguez N, Ser JD, Bennetot A, Tabik S, González AB, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges. *Information Fusion*. 2020;58:82–115. DOI: 10.1016/j.inffus.2019.12.012.
5. Lee JY, Lee Y. Integrative literature review on people analytics and implications from the perspective of human resource development. *Human Resource Development Review*. 2024;23(1):58–87.
6. Pinto GBS, Mello CE, Garcia ACB. Explainable AI in labor-market applications. In: Rocha AP, Steels L, van den Herik JH, editors. *Proceedings of the 17th International conference on agents and artificial intelligence; 2025 February 23–25; Porto, Portugal. Volume 3*. Lissabon: Science and Technology Publications; 2025. p. 1450–1457. DOI: 10.5220/0013384100003890.
7. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, editors. *Proceedings of the 31st International conference on neural information processing systems; 2017 December 4–9; Long Beach, USA. Volume 30*. New York: Curran Associates; 2017. p. 4765–4774.

8. Chaturvedi S, Chaturvedi R. Who gets the callback? Generative AI and gender bias. *arXiv:2504.21400* [Preprint]. 2025 [cited 2025 May 3]. Available from: <https://arxiv.org/abs/2504.21400>.
9. Klöpper M, Messer U. Hiding behind algorithms: people analytics and perceived fairness violation in managerial decisions [Internet]. In: Avital M, Karahanna E, Themistocleous M, editors. *Proceedings of the 32nd European conference on information systems. People first: constructing digital futures together; 2024 June 13–19; Paphos, Cyprus* [cited 2025 May 3]. Available from: https://aisel.aisnet.org/ecis2024/track07_busanalytics/track07_busanalytics/16.
10. Chen T, Guestrin C. XGBoost: a scalable tree-boosting system. In: Krishnapuram B, Shah M, editors. *KDD'16. Proceedings of the 22nd ACM SIGKDD International conference on knowledge discovery and data mining; 2016 August 13–17; San Francisco, USA*. New York: Association for Computing Machinery; 2016. p. 785–794. DOI: 10.1145/2939672.2939785.
11. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. In: Tere-desai A, Kumar V, editors. *KDD'19. Proceedings of the 25th ACM SIGKDD International conference on knowledge discovery & data mining; 2019 August 4–8; Anchorage, USA*. New York: Association for Computing Machinery; 2019. p. 2623–2631. DOI: 10.1145/3292500.3330701.
12. Spence AM. *Job market signaling*. Cambridge: Harvard University Press; 1973. 174 p.
13. Mincer J. *Schooling, experience, and earnings*. Cambridge: National Bureau of Economic Research; 1974. 152 p.
14. Гуриева СД. Феномен «тяги-выталкивай» как проявление гендерного неравенства в организационном контексте. *Russian Journal of Education and Psychology*. 2020;11(4):7–11. EDN: NWCLHP.

Статья поступила в редколлегию 19.08.2025.

Received by editorial board 19.08.2025.