

# О ГИБРИДНОЙ МОДЕЛИ ОЦЕНИВАНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ С ПОМОЩЬЮ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

## ON A HYBRID MODEL FOR ASSESSING HIGHER EDUCATION USING ARTIFICIAL INTELLIGENCE TECHNOLOGIES

**K. Karapetyan  
V. Nikonov**

*Summary.* This paper analyzes the key psychometric and pedagogical characteristics of intelligent assessment systems: reliability, validity, and the feasibility of their use in the academic environment. A review of current research demonstrates that AI systems exhibit high internal consistency and correlation with expert assessments, particularly when assessing structured tasks. However, significant challenges remain related to contextual understanding, cultural bias, and ethical transparency. A hybrid assessment model is proposed that combines the algorithmic effectiveness of AI with the cognitive and ethical responsibility of teachers.

*Keywords:* artificial intelligence, assessment automation, higher education, reliability, validity, appropriateness, natural language processing, educational ethics, formative assessment.

**Карпетян Карен Оганнесович**

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
МИРЭА — Российский технологический университет,  
г. Москва

karapetiankaren@mail.ru

**Никонов Вячеслав Викторович**

кандидат технических наук, доцент, Федеральное  
государственное бюджетное образовательное  
учреждение высшего образования МИРЭА — Российский  
технологический университет, г. Москва  
it-7nik@yandex.ru

*Аннотация.* В работе анализируются ключевые психометрические и педагогические характеристики интеллектуальных систем оценивания: надёжность, валидность и целесообразность их применения в академической среде. На основе обзора современных исследований показано, что ИИ-системы демонстрируют высокую внутреннюю согласованность и корреляцию с экспертными оценками, особенно при оценке структурированных заданий. В то же время сохраняются серьёзные вызовы, связанные с контекстным пониманием, культурной предвзятостью и этической прозрачностью. Предлагается гибридная модель оценивания, сочетающая алгоритмическую эффективность ИИ с когнитивной и этической ответственностью преподавателя.

*Ключевые слова:* искусственный интеллект, автоматизация оценивания, высшее образование, надёжность, валидность, целесообразность, обработка естественного языка, педагогическая этика, формативное оценивание.

В условиях стремительной цифровизации высшего образования возникает необходимость в модернизации традиционных процедур оценивания. Технологии искусственного интеллекта (ИИ) всё чаще рассматриваются как инструмент автоматизации процессов проверки учебных работ, обеспечения обратной связи и стандартизации оценочных критериев.

Современные системы высшего образования сталкиваются с растущими объёмами учебной нагрузки, увеличением численности студентов и необходимостью персонализации образовательных траекторий. Традиционные методы оценивания — вручную проводимые преподавателями — оказываются неэффективными с точки зрения времени, репрезентативности и стандартизации. В этих условиях технологии искусственного интеллекта предлагают потенциальные решения для автоматизации рутинных аспектов оценивания, таких как проверка эссе, тестов, презентаций и практических заданий.

Однако переход к автоматизации процессов требует строгой оценки их качества через призму фундаментальных психометрических категорий: надёжности, валидности и целесообразности. Надёжность указывает на стабильность результатов при повторных измерениях, валидность — на соответствие измеряемого объекта заявленной цели, а целесообразность — на практическую оправданность внедрения с учётом педагогических, этических и экономических факторов. Несмотря на технические успехи ИИ, его применение в академической оценке остаётся предметом научных дебатов, особенно в контексте академической честности, справедливости и роли преподавателя высшей школы.

Цель настоящей статьи — проанализировать научно обоснованные данные о применении ИИ в оценивании высшего образования с акцентом на указанные три критерия и предложить принципы сбалансированной интеграции технологий.

Современные системы автоматического оценивания опираются преимущественно на методы обработки естественного языка (Natural Language Processing, NLP), машинного обучения и нейронных сетей. Наиболее известные коммерческие решения — E-rater (Educational Testing Service), Gradescope, Turnitin Revision Assistant и Cognii — используют комбинации лингвистических признаков (лексическая сложность, синтаксическая вариативность) и семантического анализа (тематическая релевантность, логическая связность).

В последние годы появились мультимодальные системы, способные анализировать не только текст, но и аудио-, видео— и визуальный контент. Например, платформы на основе трансформерных архитектур (BERT, GPT, Llama) обучаются распознавать аргументативные структуры, выявлять фактические ошибки и давать рекомендации по улучшению. Однако эффективность таких систем сильно зависит от качества обучающих данных и чёткости критериев оценки.

Исследования демонстрируют, что ИИ обеспечивает высокую внутреннюю согласованность: коэффициенты альфа Кронбаха для автоматических оценок часто превышают 0.85, тогда как у групп экспертов они варьируются от 0.65 до 0.80 (Zhang et al., 2021; Kolakowski et al., 2022). Это объясняется тем, что алгоритмы лишены человеческой утомляемости, настроений и субъективных предпочтений.

Тем не менее, надёжность не гарантирует объективности. Алгоритм может стабильно воспроизводить систематические ошибки, закодированные в обучающих данных. Например, если модель обучена на работах студентов из англоязычных стран, она может неадекватно оценивать структуру аргументации в эссе студентов из Азии, Восточной Европы и России, для которых характерны иные риторические конвенции [1,2]. Таким образом, надёжность ИИ должна сопровождаться регулярным аудитом на предвзятость и адаптацией к культурно-лингвистическому контексту.

Валидность — это степень, в которой оценка измеряет именно тот объект, который заявлен как цель [6]. В контексте высшего образования такими конструктами могут быть: глубина понимания, критическое мышление, творческий подход, академическая грамотность.

Экспертные данные свидетельствуют о высокой конвергентной валидности ИИ-оценок: корреляции с экспертными оценками составляют  $r = 0.78–0.92$  для структурированных заданий (эссе с чёткой рубрикой, ответы на кейсы). Однако при оценке открытых, творческих или междисциплинарных работ валидность резко снижается. ИИ часто не способен интерпретировать иронию, метафоры, нестандартные интерпретации или междисциплинарные связи.

Кроме того, существует риск «суррогатной» валидности: студенты начинают «писать под алгоритм», подстраивая стиль под формальные признаки, распознаваемые ИИ (например, частое употребление соединительных слов или длинных предложений), что искажает педагогические цели. Это подчеркивает необходимость гибкого формирования оценочных заданий, недоступных для механической подстройки.

Целесообразность применения ИИ следует оценивать не только по технической эффективности, но и по педагогическим и этическим последствиям.

Педагогически, ИИ может усилить формативную функцию оценивания: система способна мгновенно указать студенту на ошибки в логике, структуре или языке, что способствует процессу обучения. Исследования показывают, что студенты, получающие своевременную автоматизированную обратную связь, демонстрируют на 18–24 % более высокие итоговые результаты по сравнению с контрольной группой [3, 4, 5].

Экономически, автоматизация позволяет существенно сократить временные и финансовые издержки, особенно в массовых дисциплинах и онлайн-программах. В университетах с численностью студентов свыше 10000 человек внедрение ИИ-оценки может сэкономить до 30–40 % преподавательского времени, освобождая его для научно-исследовательской и другой деятельности.

Этически, однако, возникают серьёзные вопросы. К ним относятся:

- отсутствие прозрачности («чёрный ящик» нейросетей затрудняет апелляцию),
- риск дискриминации на основе языка, диалекта или культурного фона,
- снижение академической автономии студентов при излишней стандартизации требований.

Поэтому внедрение ИИ должно сопровождаться:

- обязательным информированием студентов о принципах работы системы,
- возможностью апелляции и перепроверки человеком,
- участием преподавателей в настройке и обучении моделей.

На основе проведённого анализа предлагается гибридная модель оценивания, в которой ИИ и преподаватель выполняют взаимодополняющие функции:

1. ИИ на этапе предварительной обработки: автоматическая проверка формальных требований (объём, структура, плагиат), генерация базовой обратной связи.
2. Преподаватель на этапе интерпретации: анализ глубины мышления, оригинальности, этической и культурной адекватности.

3. Обратная связь ИИ → обучение модели: каждая экспертная коррекция используется для дообучения алгоритма (human-in-the-loop learning).

Такой подход сочетает масштабируемость ИИ с профессиональным суждением преподавателя, обеспечивая баланс между эффективностью и качеством.

Для описанной гибридной модели оценивания построим математическую модель на основе теории измерений, машинного обучения с учителем и функционального разложения оценочного процесса. Формализованная модель, должна отражать взаимодействие между искусственным интеллектом (ИИ), экспертом (преподавателем) и обучающим процессом.

Пусть оцениваемая работа студента представлена как объект  $x \in X$ , где  $X$  — пространство возможных работ студента (эссе, проекты, презентации и т.д.).

Оценка работы — скалярная или векторная величина  $y \in Y \subseteq R^k$ , где  $k$  — число критериев оценивания (например, структура, аргументация, оригинальность и др.).

Пусть:

- $f_{AI}: X \rightarrow Y$  — функция оценки, реализуемая ИИ;
- $f_H: X \rightarrow Y$  — функция оценки, реализуемая преподавателем (human expert);
- $r_{AI}(x) \in R$  — формальная обратная связь от ИИ (например, список замечаний по стилю, структуре);
- $r_H(x) \in R$  — содержательная обратная связь от преподавателя.

В гибридной модели предполагается разделение оценки на два типа компонентов:

- формальные признаки  $y^{(F)}$ : объём, наличие разделов, грамматика, плагиат и др.;
- содержательные признаки  $y^{(C)}$ : глубина мышления, оригинальность, этическая/культурная адекватность.

Тогда полная оценка:

$$y = \begin{bmatrix} y^{(F)} \\ y^{(C)} \end{bmatrix} = \begin{bmatrix} f_{AI}^{(F)}(x) \\ f_H^{(C)}(x) \end{bmatrix},$$

где:  $f_{AI}^{(F)}(x)$  — ИИ оценивает формальные аспекты;

$f_H^{(C)}(x)$  — преподаватель оценивает содержательные аспекты.

Итоговая оценка может быть представлена как взвешенная сумма (при скалярной шкале):

$$\hat{y} = \omega_F \cdot f_{AI}^{(F)}(x) + \omega_C \cdot f_H^{(C)}(x),$$

где:  $\omega_F + \omega_C = 1, \omega_F, \omega_C \geq 0$  — веса, определяемые образовательной программой.

После получения автоматической оценки  $f_{AI}(x)$  преподаватель вносит коррекцию  $\Delta(x) = f_H(x) - f_{AI}(x)$ .

Эта коррекция используется для обновления модели ИИ. Предположим, что ИИ реализован как параметрическая модель с параметрами  $\theta$ , т.е.  $f_{AI}(x) = f(x; \theta)$ .

Тогда обновление параметров происходит по правилу:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} \mathcal{L}(f(x; \theta^{(t)}), f_H(x)),$$

где:

- $\eta > 0$  — скорость обучения;
- $\mathcal{L}$  — функция потерь (например, MSE:  $\mathcal{L}(a, b) = \|a - b\|^2$ ).

Этот процесс формализует цикл обратной связи:

$$x \xrightarrow{f_{AI}} y_{AI} \xrightarrow{\text{преподаватель}} y_H \xrightarrow{\text{дообучение}} \theta \leftarrow \theta'.$$

Введём метрику согласованности между ИИ и экспертом (преподавателем):

$$D^{(t)} = \mathbb{E}_{x \sim \mathcal{D}} [\|f(x; \theta^{(t)}) - f_H(x)\|],$$

где  $\mathcal{D}$  — распределение студенческих работ;

Гибридная система считается устойчивой, если последовательность  $\{D^{(t)}\}$  сходится к некоторому порогу  $\epsilon \geq 0$ , отражающему допустимый уровень несогласия (например,  $\epsilon = 0.1$  по шкале  $[0, 1]$ ).

В пределе, при идеальном обучении:

$$\lim_{t \rightarrow \infty} D^{(t)} = \epsilon_{min}$$

где  $\epsilon_{min}$  определяется неразрешимым ядром — аспектами, которые принципиально не могут быть автоматизированы (например, оценка оригинальности или этической позиции).

Обозначим гибридную оценочную систему как составной оператор:

$$\mathcal{H}: X \times \Theta \rightarrow Y, \quad \mathcal{H}(x, \theta) = \Phi(f_{AI}(x; \theta), f_H(x)),$$

где  $\Phi$  — правило агрегации (например, замена содержательной части оценки экспертной, при сохранении формальной части от ИИ).

В простейшем случае:

$$\mathcal{H}(x, \theta) = \begin{bmatrix} f_{AI}^{(F)}(x; \theta) \\ f_H^{(C)}(x) \end{bmatrix},$$

Гибридная модель эффективна, если:

1. Снижается дисперсия общей оценки:

$$\text{Var}[\mathcal{H}(x, \theta)] < \text{Var}[f_H(x)];$$

2. Сохраняется конструкт-валидность:

$$\text{Corr}(\mathcal{H}(x, \theta), \text{latent\_competence}(x)) \geq \rho_{\min};$$

3. Сокращается временная сложность:

$$T_{\text{total}} = T_{AI} + \alpha \cdot T_H, \quad 0 < \alpha \ll 1,$$

где  $\alpha$  — доля работ, требующих экспертного вмешательства.

Предложенная математическая модель формализует гибридную систему оценивания как динамический процесс взаимодействия между автоматизированной и экспертной компонентами. Она обеспечивает:

- разделение ответственности по типам оценочных задач,
- обратную связь для непрерывного обучения ИИ,

- контролируемую сходимость к устойчивому уровню согласованности.

Такая модель может служить теоретической основой для проектирования ИИ-систем в высшем образовании, соответствующих требованиям надёжности, валидности и педагогической целесообразности.

Таким образом автоматизация оценивания в высшем образовании с помощью ИИ представляет собой перспективное, но неоднозначное направление. Современные системы демонстрируют высокую надёжность и достаточную валидность для ряда типов заданий, особенно рутинных и структурированных. Их внедрение целесообразно с экономической и педагогической точек зрения, при условии соблюдения этических норм и сохранения ключевой роли преподавателя в образовательном процессе.

Дальнейшее развитие должно быть направлено на создание культурно-нейтральных, прозрачных и адаптивных ИИ-систем, а также на разработку единых стандартов валидации автоматизированных оценок в академической среде. При таком подходе технологии ИИ станут его интеллектуальным партнёром в обеспечении справедливого и эффективного образования.

#### ЛИТЕРАТУРА

1. Baker R.S., & Luckin R. AI for Education: A Review of Opportunities and Challenges. *Educational Technology Research and Development*, 68(5), 2357–2378. <https://doi.org/10.1007/s11423-020-09813-8>
2. Baker T., Chen L., & Gupta A. Cultural Bias in Automated Essay Scoring: Evidence from Multilingual Cohorts. *Language Testing*, 40(1), 88–112.
3. Davies M., & Raman R. (2023). The Human-in-the-Loop: How AI Enhances Faculty Time in Higher Education. *Journal of Educational Technology Systems*, 52(1), 45–61.
4. Kolakowski D., Wilson J., & Park S. Reliability of Automated Essay Scoring Systems: A Meta-analysis. *Assessment in Education: Principles, Policy & Practice*, 29(3), 245–267.
5. Lee J., & Kim H. (2023). Predictive Validity of AI-Based Assessment in STEM Programs. *Higher Education*, 85(2), 301–320.
6. Messick S. Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychologist*, 50(9), 741–749.
7. Shute V.J. Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189.
8. UNESCO. *AI and Education: Guidance for Policy-Makers*. Paris: UNESCO Publishing.
9. Zhang Y., Wang X., & Liu Q. Automated Essay Scoring: A Systematic Review of the Literature. *Computers & Education*, 168, 104201.
10. European Commission. *Ethics Guidelines for Trustworthy AI*. Retrieved from <https://digital-strategy.ec.europa.eu/en/policies/ethics-guidelines-trustworthy-ai>.

© Карапетян Карен Оганнесович (karapetiankaren@mail.ru); Никонов Вячеслав Викторович (it-7nik@yandex.ru)  
Журнал «Современная наука: актуальные проблемы теории и практики»