

ЭКСПЕРТНАЯ ОЦЕНКА ТОЧНОСТИ РАБОТЫ НЕЙРОСЕТИ ДЛЯ ОЦЕНКИ ПРАВИЛЬНОСТИ РЕШЕНИЯ ЗАДАНИЙ ШКОЛЬНОЙ ПРОГРАММЫ

А. А. Артамонов, Т. Ц. Намдаг, Е. М. Косарева

*Белорусский государственный университет информатики
и радиоэлектроники, г. Минск*

Представлен протокол экспертной оценки, направленный на определение точности ответов нейросетевой модели, используемой для проверки решения заданий школьного уровня.

Ключевые слова: искусственный интеллект, образовательные технологии, оценка точности, экспертная валидация, протокол тестирования.

EXPERT EVALUATION OF THE ACCURACY OF A NEURAL NETWORK FOR ASSESSING THE CORRECTNESS OF SOLVING SCHOOL CURRICULUM ASSIGNMENTS

A. A. Artamonov, T. Ts. Namdag, E. M. Kosareva

Belarusian state university of informatics and radioelectronics, Minsk

The article presents an expert evaluation protocol aimed at determining the accuracy of responses from a neural network model used to check the solutions to school-level assignments.

Keywords: user interface, medical diagnostic systems, usability engineering.

Актуальность разработки данного протокола обусловлена необходимостью объективной проверки качества работы интеллектуальных систем, применяемых в образовательной среде. В условиях активного внедрения технологий искусственного интеллекта в учебный процесс возрастает значение системной и воспроизводимой методики оценки точности генерируемых ответов.

Типы заданий, используемых в качестве выборки для оценки точности ответов нейросети, можно классифицировать следующим образом:

1. Тестовые задания (с выбором ответа) – предполагают выбор одного или нескольких правильных вариантов из предложенных [4].
2. Краткие ответы – требуют воспроизведения конкретного факта, формулы или определения. Основным критерий – фактическая точность [4].
3. Развернутые ответы (эссе, объяснения, решения задач) – оцениваются по полноте раскрытия, логической структуре и языковой корректности [1].
4. Задания на вычисления – предполагают получение численного результата. Оценка проводится через сравнение с эталоном с допустимой погрешностью [4].

Оценка точности ответов нейросети производится на основании совокупности факторов, отражающих как содержательные, так и формальные характеристики решений.

Каждый ответ нейросети оценивается по следующим критериям:

- фактическая точность;
- логическая связность;
- полнота раскрытия;
- языковая корректность;
- соответствие уровню задания [3].

Метод оценки точности ответов нейросети основан на сопоставлении с эталонными решениями и последующей экспертной валидации.

Для каждого критерия выставляется балл от 0 до 2:

- 0 – несоответствие критерию;
- 1 – частичное выполнение;
- 2 – полное выполнение.

После этого рассчитывается интегральный показатель точности представленные по формуле

$$T = \frac{\sum B_i}{2n},$$

где B_i – сумма баллов по критериям для i -го задания; n – число заданий.

Полученный процент отражает степень совпадения ответов нейросети с эталоном и позволяет классифицировать уровень ее точности. Модель демонстрирует соответствие учебным стандартам при точности ответов 85–100 %.

Результаты оценки точности работы модели представлены в таблице.

Результаты оценки точности работы модели

Название предмета	Тип задания	Процент правильности ответа
Математика	Тестовое задание	100
Математика	Вычислительное задание	100
Математика	Краткий ответ	100
Математика	Развернутый ответ	96

Разработанная методика была применена для оценки ответов нейросети, обученной на базе школьных учебных данных. В процессе тестирования были проанализированы задания различных типов. На основании экспертных оценок и вычисления интегрального показателя установлено, что нейросетевая модель продемонстрировала высокий уровень точности. Особенно высокий результат был зафиксирован при решении вычислительных и тестовых заданий, где степень совпадения с эталоном достигала 100 %. Незначительные отклонения отмечены преимущественно в развернутых ответах, требующих интерпретации и логического обоснования.

Таким образом, предложенная система экспертной оценки доказала свою эффективность как инструмент объективного анализа когнитивных возможностей нейросетевых моделей. Полученные результаты подтверждают, что разработанная нейросетевая модель демонстрирует устойчивую способность к воспроизведению корректных решений в рамках школьной программы и может быть использована в образовательных целях. Представленный протокол может служить основой для дальнейшей стандартизации процедур оценки искусственного интеллекта в сфере образования и внедрения адаптивных систем контроля качества обучающих моделей.

Литература

1. Системы искусственного интеллекта в образовании: методики тестирования и валидации – URL: <https://cyberleninka.ru/article/n/sistemy-iskusstvennogo-intellekta-v-obrazovanii/viewer> (дата обращения: 14.10.2025).
2. Методы оценки эффективности интеллектуальных систем. – URL: <https://elibrary.ru/item.asp?id=48563291> (дата обращения: 14.10.2025).
3. ГОСТ Р 57197-2016. Системы искусственного интеллекта. Термины и определения. – М. : Стандартинформ, 2016. – 20 с.