Туревская Елена Ильинична
канд. психол. наук, доцент
ФГБОУ ВО «Тульский государственный педагогический
университет им. Л.Н. Толстого»
г. Тула, Тульская область

МОДИФИКАЦИЯ ОЦЕНОЧНЫХ СРЕДСТВ В УСЛОВИЯХ КОНКУРЕНЦИИ С НЕЙРОСЕТЯМИ: РИСК ИЛИ ГАРАНТИЯ УСПЕХА

Аннотация: *статья посвящена проблеме снижения диагностической функции традиционных фондов оценочных средств в педагогическом вузе вследствие массового использования студентами нейросетей. Автор констатируют, что устоявшиеся форматы контроля (эссе, рефераты, тесты, решение кейсов) перестали оценивать уровень сформированности профессиональных компетенций, фактически измеряя лишь навыки промпт-инжиниринга студентов. В качестве решения предлагается концептуальный сдвиг: перенос фокуса оценивания с легко генерируемого текстового результата на процессуальную составляющую обучения и верификацию знаний. Автор подчеркивает, что ИИ не должен запрещаться или игнорироваться, а должен интегрироваться в учебный процесс как инструмент, работу которого студент обязан критически проанализировать.*

Ключевые слова: *оценочные средства, искусственный интеллект (ИИ), нейросеть, промпт-инжиниринг.*

В последние годы приходится констатировать тот факт, что наш фонд оценочных средств, который мы с коллегами кропотливо выстраивали более десятилетия, опираясь на методические рекомендации ведущих вузов, фактически обесценился.

Мы столкнулись с парадоксом: наши проверенные временем инструменты – эссе, рефераты, аналитические справки, решение кейсов – перестали оценивать профессиональные знания и компетенции будущего педагога. Они начали оценивать навык промпт-инжиниринга (создание инструкций для искус-

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

ственного интеллекта (ИИ), чтобы получать от него максимально точные, предсказуемые и полезные результаты). Мы получаем от студентов безупречно структурированный, терминологически грамотный текст и понимаем: мы только что оценили не профессиональную компетенцию студента, а его умение составить запрос так, чтобы нейросеть не «галлюцинировала».

Согласно методическим рекомендациям Института качества высшего образования НИТУ МИСИС, оценочные средства – это контрольные задания, описание форм и процедур, предназначенных для определения уровня сформированности необходимых компетенций [1].

Фонд оценочных средств – контрольно-измерительных материалов, предназначенных для оценивания компетенций на разных стадиях обучения студентов, а также для аттестационных испытаний выпускников на соответствие (или несоответствие) уровня их подготовки требованиям ФГОС по направлению подготовки или специальности.

В процессе подготовки рабочих программ для Педагогического направления подготовки, был подобран фонд оценочных средств диагностики сформированности каждой компетенции, представленной в программе. Ежегодно проводилась работа по коррекции оценочных средств, приведению их в соответствие с изменениями в профессиональных стандартах.

Традиционно профессиональные компетенции оценивается по результатам выполнения работы на практических занятиях (сообщения), тестовых заданий; составления психологического словаря, Maind-map; разработки диагностической программы, интерпретации результатов тестирования, написания эссе, решения кейсов, контрольных работ, сдачи зачета и экзамена.

Однако, в современных условиях, из всех средств массовой информации, включая телевидение, соц. сети и мессенджеры, мы узнаем о разных нейросетях (ChatGPT, GigaChat, BeeGPT и т. п.), которые направлены именно на студентов и позиционируются как «помощники» для выполнения контрольных заданий и успешной сдачи сессии. Появляется вопрос валидности нашей оценки, мы не можем быть уверены, что используя традиционные оценочные средства, прове-

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

ряем сформированные компетенции, умение мыслить, а не качество и эффективность работы нейросети.

ГОСТ 2021 года трактует искусственный интеллект (ИИ) как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (в том числе самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека» [2]. Нейросеть рассматривается нами как часть ИИ, один из конкретных методов внутри искусственного интеллекта.

Использование нейросетей несет в себе следующие риски.

1. У студента появляется ложное впечатление своей компетентности в определенном вопросе. Студент сдает идеально структурированную, терминологически грамотную работу, но на защите выясняется, что он не может объяснить, как он пришел к таким выводам. В итоге оценивается работа нейросети, а не мышление студента.

2. Снижение долговременной памяти. При постоянной передаче функции поиска информации, ее анализа и синтеза нейросетям, у студентов не формируются долгосрочные нейронные связи. Знание не переходит из оперативной памяти в долговременную. Студент знает, где быстро найти ответ, но не знает сам ответ, что делает невозможным его применение в стрессовой ситуации или при отсутствии интернета.

3. Потеря авторского стиля. Тексты, сгенерированные ИИ, имеют узнаваемый «нейтральный» стиль, усредненную лексику и отсутствие индивидуальных речевых ошибок (авторский стиль). Это обесценивает уникальность мышления – все начинают писать одинаковыми, гладкими фразами.

4. Галлюцинации и «фабрикация» фактов. ИИ склонен выдумывать источники, статистические данные и цитаты, которые выглядят крайне правдоподобно. Студент, не обладающий глубокой эрудицией в предмете, не способен распознать эту ложь, транслируя ее в своей работе как истину. Это дискредитирует научную базу исследования (см. Пример 2).

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

5. Снижение мотивации к обучению. Если высокие баллы можно получить за две минуты генерации текста с помощью ИИ, у студента пропадает потребность тратить время на самостоятельное чтение первоисточников, конспектирование и самостоятельный анализ прочитанного.

Пример 1. Задание студентам: напишите психологическую характеристику ребенку младшего школьного возраста. При написании характеристики укажите методики диагностики (название методики, авторы, на какие психические процессы направлена).

Студент дает задание нейросети, присылает преподавателю ответ.

Ответ нейросети.

Ребенок младшего школьного возраста (6–9 лет) находится в периоде активного развития своих психических процессов. На данном этапе развития происходит значительное изменение в познавательной, эмоциональной и социальной сферах.

Для диагностики перечисленных психических процессов у ребенка младшего школьного возраста могут использоваться следующие методики.

1. Тесты на развитие логического мышления (например, тесты решения логических задач А. Н. Леонтьева).

2. Тесты на оценку речевых навыков, слушания и понимания речи (например, тесты на слуховое восприятие А. С. Белых).

3. Тесты на определение уровня готовности к чтению и письму (например, тесты на анализ звуков речи и определение букв А. Н. Халатовой).

4. Методики на определение уровня эмоциональной саморегуляции (например, методики на определение уровня осознанности эмоций Р. И. Багмаска).

Ответ, который сдал студент, воспользовавшись нейросетью, соответствует заданию, написан научным языком, но содержит множество артефактов (в психологической характеристике написаны общие слова, подходящие не только к заданному возрастному периоду, ни одной названной методики с названной целью и автором не существует).

Пример 2.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

Задание студентам: подберите диагностическую методику мотивации учения для старших дошкольников. В описании методики укажите: название, автора, цель, краткое содержание (что оценивает). Аргументируйте свой ответ.

Ответ нейросети.

Одной из эффективных диагностических методик мотивации учения для старших дошкольников может быть методика «Сонячные круги» Ангелины Серафимовой. Цель этой методики – выявить интересы, мотивы и потребности детей в процессе обучения.

Суть методики заключается в том, что каждому ребенку предлагается нарисовать «солнечное лицо» на листе бумаги. Каждый «лучик» солнца представляет изучаемый вид деятельности или широкую сферу интересов ребенка (например, игру, рисование, рукоделие, чтение книг, музыку, спорт и т. д.). Ребенку предлагается раскрасить «лучи» солнца в разные цвета в зависимости от степени интереса к конкретной деятельности.

Анализируя итоговую картинку с «солнечным лицом», педагог может сделать выводы о предпочтениях ребенка, его мотивации к разным видам деятельности, а также о его ценностях и потребностях. Например, если большая часть «лучиков» окрашена в яркий цвет, это означает, что ребенок проявляет большой интерес к определенной сфере. Если же часть «лучиков» остается неокрашенной или слабо окрашена, это может говорить о недостаточной мотивации ребенка к конкретной деятельности.

Данная методика помогает педагогу индивидуализировать образовательный процесс, учитывать индивидуальные особенности и мотивацию каждого ребенка. Анализируя результаты методики «Сонячные круги», педагог может адаптировать содержание обучения, предлагая детям разнообразные виды деятельности, которые соответствуют их интересам и мотивам. Таким образом, методика помогает создать поддерживающую среду, способствующую развитию у детей позитивной мотивации к учению и саморазвитию.

Преподаватель может использовать нейросеть для анализа данного ответа.

Таблица 1

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

Выявление артефактов (искажений и манипуляций)

Тип артефакта	Цитата из текста	Суть проблемы
Анонимный авторитет (и отсутствие источника)	«...методика «Сонячные круги» Ангелины Серафимовой»	В тексте приводится имя автора, однако оно не подкреплено ссылками на научные публикации, методические пособия или реестры психологических методик. В академической и педагогической среде такая методика не является общепризнанной или стандартизированной. Это создает иллюзию экспертности там, где может быть частный авторский прием
Отсутствие психометрической проверки	«Одной из эффективных диагностических методик...»	Слово «диагностическая» подразумевает наличие данных о валидности (измеряет ли методика то, что должна) и надежности (будет ли результат одинаковым при повторном тестировании). В тексте нет никаких ссылок на результаты апробации, стандартизации или психометрической проверки
Нарушение причинно-следственной связи (Ложная корреляция)	«Если большая часть «лучиков» окрашена в яркий цвет, это означает, что ребенок проявляет большой интерес...»	Цвет в психологии не имеет универсальной диагностической силы. Ребенок может выбрать яркий цвет просто потому, что он перед ним лежит, или потому, что этот карандаш самый острый. Яркость цвета не является валидным индикатором «интереса» или «ценностей» без дополнительной калибровки и контрольных вопросов
Подмена тезиса (диагностика вместо проекции)	«Цель этой методики – выявить интересы, мотивы и потребности... а также о его ценностях и потребностях»	Рисуночные методики являются проективными. Они выявляют не объективные «мотивы и потребности», а лишь субъективное представление ребенка о них в данный момент времени. Текст же преподносит результат рисования как объективный факт (диагноз), что является грубым искажением сути проективного тестирования.
Гипербола и необоснованное обобщение	«...помогает педагогу индивидуализировать образовательный процесс...»	Один рисунок не дает данных для полноценной индивидуализации всего образовательного процесса. Утверждение приписывает частному, поверхностному инструменту системообразующую функцию, которой он не обладает

Оценка степени достоверности: 25% (Недостоверно).

Обоснование: текст представляет собой описание частного, невалидизированного педагогического приема, который выдается за научно обоснованную «диагностическую методику».

1. *Фактологическая недостоверность.* Утверждение о том, что методика является «эффективной диагностической», ложно без данных о ее валидности и надежности. Ссылка на «Ангелину Серафимову» без указания источника (статьи, книги, года издания) делает информацию непроверяемой.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

2. *Логические искажения.* Текст строится на ложной корреляции между цветом карандаша и глубиной мотивационной сферы ребенка. Игнорируется тот факт, что выбор цвета детьми дошкольного возраста часто обусловлен случайными факторами, а не сложной внутренней мотивацией.

3. *Манипулятивность.* Текст использует педагогическую тревожность (страх не учесть индивидуальность ребенка) и предлагает простое, визуально приятное, но научно несостоятельное решение. Использование термина «диагностическая» по отношению к проективному рисованию без строгих рамок интерпретации является прямым введением в заблуждение.

Итоговое заключение нейросети.

Методика «Сонячные круги» в описанном виде не обладает научной достоверностью: она не позволяет объективно измерить мотивы или потребности ребенка, так как связь между цветом раскрашенного луча и реальной мотивацией не доказана, его результаты категорически нельзя использовать для «индивидуализации образовательного процесса» или постановки каких-либо педагогических «диагнозов».

Мы не можем запретить использование ИИ, но и игнорировать факт, что старые оценочные средства не выполняют свою функцию, невозможно. Возможный вариант очередной модернизации оценочных средств в условиях конкуренции с ИИ – перенести фокус оценивания с результата (текста, который легко сгенерировать с помощью ИИ) на процесс выполнения задания (как студент пришел к выводу) и проверку (способен ли студент подтвердить каждый вывод своей работы здесь и сейчас без доступа к сети) [3].

При создании модернизированных оценочных средств следует придерживаться следующих принципов:

− привязка к уникальному контексту: использование данных, которых нет в обобщенных данных ИИ (личный опыт студента, наблюдение за конкретным человеком здесь и сейчас);

− требование метакогниции: студент должен описывать не только что он сделал, но и как он думал в процессе.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

− физическое присутствие и асинхронность: выполнение задания в аудитории без гаджетов или с обязательной устной защитой;

− синтез разномодальных данных: объединение текста, графики, аудио и личного взаимодействия, что сложно передать одной текстовой командой.

ИИ должен стать в учебном процессе не способом обхода оценки, а инструментом, работу которого студент должен критически проанализировать и описать. Такой подход рассматривается как единственный способ сохранить валидность оценки и гарантировать соответствие подготовки выпускников требованиям ФГОС в условиях конкуренции с нейросетями.

Список литературы

1. Туревская, Е. И. Использование оценочных средств в процессе формирования компетенций у бакалавров непедагогических направлений подготовки / Е. И. Туревская, Н. А. Хвалина // Обеспечение качества учебного процесса: традиции и инновации : материалы XLII учеб.-метод. конф. профессорско-преподавательского состава, аспирантов, магистрантов, соискателей ТГПУ им. Л.Н. Толстого (Тула, 15 апреля 2015 г.). – Тула : Тульский государственный педагогический университет им. Л.Н. Толстого, 2015. – С. 269–272.

2. ГОСТ Р 59895–2021 Технологии искусственного интеллекта в образовании. Общие положения и терминология : введ. 26.11.2021. – Москва : Федеральное агентство по техническому регулированию и метрологии, ФГАО УВО «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ), 2021. – 12 с. – URL: https://metall.world/gostsfile/24699 (дата обращения: 19.11.2026). – Текст : электронный.

3. Воробьева, М. В. Применение нейросетей в образовательном процессе: примеры, возможности и риски / М. В. Воробьева // Педагогическое образование и наука. – 2024. – № 3.

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)