

П. В. Мучин^{1✉}, М. П. Мучин²

Искусственный интеллект в образовании: сравнительное тестирование ChatGPT 4o, DeepSeek-R1 и Grok 3 по направлению подготовки «Техносферная безопасность»

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация

² Сибирский государственный университет телекоммуникаций и информатики, г. Новосибирск, Российская Федерация
e-mail: p.v.muchin@ssga.ru

Аннотация. В работе представлен следующий этап исследования возможности применения искусственного интеллекта в учебном процессе при подготовке специалистов техносферной безопасности. Отдельные вопросы (всего 25) направления подготовки студентов техносферной безопасности, предложенные платной версии ChatGPT-4 в 2023 году, были в 2025 году вновь предложены бесплатным версиям ИИ: ChatGPT 4o, DeepSeek-R1 и Grok 3. Дополнительно 5 вопросов из 25 протестированы платной версией ИИ ChatGPT o3-mini-high. Выполнен сравнительный анализ полученных ответов и подготовлены рекомендации о возможности и целесообразности использования ИИ в учебном процессе.

Ключевые слова: нейронные сети, искусственный интеллект, ChatGPT-4, ChatGPT 4o, DeepSeek-R1, Grok 3, ChatGPT o3-mini-high, обучение, техносферная безопасность

P. V. Muchin^{1✉}, M. P. Muchin²

AI in Education: Comparative Testing of ChatGPT 4o, DeepSeek-R1 and Grok 3 in the Field of Study «Technosphere Security»

¹Siberian state university of Geosystems and Technologies,
Novosibirsk, Russian Federation

²Siberian state university of Telecommunications and Informatics,
Novosibirsk, Russian Federation
e-mail: p.v.muchin@ssga.ru

Annotation. The paper presents the next stage of the study of the possibility of using artificial intelligence in the educational process in the training of technosphere safety specialists. Individual questions (25 in total) of the technosphere safety student training area, proposed to the paid version of ChatGPT-4 in 2023, were again proposed to the free versions of AI in 2025: ChatGPT 4o, DeepSeek-R1 and Grok 3. Additionally, 5 questions out of 25 were tested by the paid version of the ChatGPT o3-mini-high AI. A comparative analysis of the received answers was performed and recommendations were prepared on the possibility and feasibility of using AI in the educational process.

Keywords: neural networks, artificial intelligence, ChatGPT-4, ChatGPT 4o, DeepSeek-R1, Grok 3, ChatGPT o3-mini-high, training, technosphere security

Введение

В настоящее время в свободном доступе интернет можно увидеть поразительные результаты применения разработанных ИТ-продуктов на основе нейронных сетей. Есть примеры практически полноценной замены человека

(специалиста) в отдельных сферах деятельности информационными технологиями (программами) на основе искусственного интеллекта.

Авторами в 2023 году был проведен первый этап исследований о возможностях искусственного интеллекта в решении вопросов из разных областей техносферной безопасности. Для тестирования была выбрана стремительно набирающей популярность платная версия нейронной сети ChatGPT-4. Результаты тестирования были представлены на XIX Международном научном конгрессе «Интерэкспо ГЕО-Сибирь» в 2023 году [1].

В предыдущем исследовании [1] нами были представлены основные этапы развития искусственного интеллекта, возможности его применения, перспективы развития и др. Соответственно теоретические аспекты, связанные с развитием и применением искусственного интеллекта в настоящей работе, будут отсутствовать.

Методы и материалы

В предыдущих исследованиях (2023 г.) для платной нейронной сети ChatGPT-4 нами были составлены вопросы по отдельным направлениям подготовки бакалавров техносферной безопасности. Учитывая чаще «восторженные» отзывы о перспективах стремительно развивающихся систем ИИ, в настоящей работе продолжены исследования популярных нейронных сетей спустя два года развития по тем же вопросам техносферной безопасности [1].

В 2023 году авторами однозначно было определено, что возможности использования (даже платной версии) нейронной сети ChatGPT при подготовке специалистов техносферной безопасности в СГУГиТ, крайне неудовлетворительные. На большинство тестовых вопросов нейронная сеть не сгенерировала правильные ответы.

Студентам при освоении модулей дисциплин техносферной безопасности ChatGPT может помочь:

- при выполнении расчётных задач;
- при перефразировании текста и при необходимости повышения оригинальности текста.

Результаты

Далее в таблице 1 мы представим отдельные результаты сравнительного тестирования популярных нейронных сетей по вопросам техносферной безопасности, предложенным в 2023 году платной версии ChatGPT-4.

Таблица 1

Результаты сравнительного тестирования популярных бесплатных нейронных сетей по вопросам отдельных дисциплин подготовки студентов направления «Техносферная безопасность»

№ вопроса	Точность и полнота ответов нейронных сетей, %			
	«ChatGPT-4» (2023 г.)	«ChatGPT 4o»	«DeepSeek-R1»	«Grok 3»
1.1	10	0	100	100
1.2	10	0	70	0
1.3	5	0	5	5
1.4	100	0	100	100
1.5	5	0	10	5
2.1	70	50	80	90
2.2	5	0	80	0
2.3	30	10	80	90
2.4	50	0	0	100
2.5	30	10	80	40
3.1	0	0	80	100
3.2	20	20	10	20
3.3	60	30	80	100
3.4	40	10	80	80
3.5	60	30	100	30
4.1	0	10	100	0
4.2	90	20	100	100
4.3	10	10	100	10
4.4	100	100	100	100
4.5	100	100	100	100
5.1	60	30	80	80
5.2	40	20	50	50
5.3	40	20	80	100
5.4	100	80	100	0
5.5	80	80	100	80
Средний показатель	44,6	25,2	74,6	59,2

Примечания по ответам «ChatGPT 4o».

1. В качестве примера покажем, как «интенсивно» нейронная сеть генерирует («придумывает») несуществующие НПА. Только по трем вопросам (1.1, 1.2 и 1.3) сгенерировано 10 несуществующих НПА:

– Постановление Правительства Российской Федерации от 13 декабря 2007 года № 884 «О расследовании и учете несчастных случаев на производстве»;

– Постановление Правительства РФ от 24 июня 2003 года № 36 «Об утверждении Правил охраны труда при организации учебного процесса в образовательных учреждениях»;

– Постановление Минтруда России от 16 мая 2005 года № 30 «Об утверждении типов, размеров и состава организаций, в которых должно быть организовано управление охраной труда, а также штатное расписание и численность специалистов по охране труда»;

– Приказ Минтруда России от 5 декабря 2007 года № 677н «Об утверждении Порядка разработки и внедрения системы управления охраной труда»;

– Постановление Правительства Российской Федерации от 24 июня 2003 года № 36 «Об утверждении Правил охраны труда при организации учебного процесса в образовательных учреждениях»;

– Постановление Правительства Российской Федерации от 28 декабря 2010 года № 1226 «О требованиях охраны труда для работников»;

– ГОСТ Р 12.0.008-2019 «Системы управления охраной труда. Общие положения»;

– Правила охраны труда при выполнении строительных и монтажных работ (ПП РФ № 314 от 30 мая 2011 года);

– СанПиН 2.2.2./2.4.2620-10 «Гигиенические требования к условиям труда и охране труда при проведении строительных работ»;

– Трудовой кодекс РФ меняет произвольно, например, статья 219 ТК РФ содержит «Обучение по охране труда», а нейронная сеть генерирует ответ, что Статья 219 ТК РФ содержит обязанность назначения специалиста по охране труда, который участвует в создании СУОТ на предприятии.

2. Оценка правильности ответов «ChatGPT 4o» снижалась авторами (экспертами) в сравнении с ответами «ChatGPT-4», так как системой постоянно генерировались «придумывались» несуществующие НПА.

3. Нейронная сеть «ChatGPT 4o» не «заметила» преднамеренную ошибку в вопросе 2.4.

Примечания по ответам «DeepSeek-R1».

1. Редко, но генерирует несуществующие НПА.

2. Иногда показывает отмененные НПА.

3. Практически верно ответила на все вопросы по экологии. Предлагаются «осмысленные рассуждения».

4. Не «заметила» преднамеренную ошибку в вопросе 2.4.

Примечания по ответам «Grok 3».

1. Предлагает пространственное и максимально подробное обоснование ответов, что чаще «запутывает» пользователя.

2. СУОТ расшифровала как «служба охраны труда» и соответственно ответ предложен неверный.

3. Подчеркивает, что актуальность указанных НПА соответствует дате обращения. Тем не менее использует и отмененные НПА.

4. Достаточно часто неправильно понимает задаваемый вопрос, что

приводит к ошибкам в ответах.

Анализируя ответы тестируемых нейронных сетей, можно отметить:

– платная система «ChatGPT-4», несмотря на двухгодичное «отставание» в развитии, существенно превосходит бесплатную версию «ChatGPT 4o»: показатели правильности ответов систем – 44,6 и 25,2;

– нейронная система «DeepSeek-R1» показала лучший результат в сравнении с «Grok 3»: показатели правильности ответов систем – 74,6 и 59,2.

Для более объективной оценки платных и бесплатных версий нейронных сетей нами протестирована платная версия «ChatGPT o3-mini-high» по отдельным вопросам. Результаты представлены ниже в таблице 2.

Таблица 2

Пробное тестирование платной версии нейронной сети «ChatGPT o3-mini-high» по отдельным вопросам из подготовленного общего перечня вопросов технологической безопасности (1.4; 2.4; 3.3; 4.5; 5.3)

№ вопроса	Оценка и комментарий ответов нейронной сети «ChatGPT o3-mini-high»
1.4	Ответ верный. Дополнительно сравнила с данными статистики РФ. Сделала выводы и предложила рекомендации.
2.4	Преднамеренную ошибку в вопросе не «увидела». Отвечала на «правильную» часть вопроса. Давала ссылки на сайты.
3.3	Ответ верный. Приведены НПА и даны ссылки на методические рекомендации.
4.5	Ответ верный. Приведены ссылки на прогнозы Правительства РФ.
5.3	Ответ не полный – точность и полнота ответа примерно 80 %. Есть ссылки на НПА и сайты.

Наша оценка нейронной системы «ChatGPT o3-mini-high» однозначна – возможности платной системы существенно выше.

Заключение

По выполненным исследованиям можно сделать выводы.

1. Несмотря на двухгодичное развитие нейронных сетей (сравнение с результатами исследований 2023 года), они продолжают показывать крайне низкие результаты по вопросам подготовки специалистов направления «Техносферная безопасность».

2. Если оценивать по пятибалльной системе, то можно представить следующие оценки ответам нейронных сетей:

- «ChatGPT-4» – 2+;
- «ChatGPT 4o» – 2-;
- «DeepSeek-R1» – 4-;

- «Grok 3» – 3;
- «ChatGPT o3-mini-high» – 4-.

3. «ChatGPT o3-mini-high» пропустила ошибку в вопросе, что снизило ответ на 1 балл. По нашему мнению, при большем количестве вопросов, скорее всего, эта платная нейронная сеть покажет высокий результат.

4. Нейронные сети могут быстро изменять текст, тем самым позволяя «нерадивым» студентам выдавать чужие работы как свои.

5. При специальном задании, нейронные сети могут существенно повышать оригинальность текста через перефразирование, но при этом возникает угроза «диагностирования» системами «Антиплагиат» генерации текста.

6. Оценивая полученные результаты необходимо учитывать, что тестирование выполнялось в русскоязычном сегменте интернет, а вопросы затрагивали меняющееся правовое поле в области безопасности.

7. На наш взгляд, наибольшая проблема применения бесплатных нейронных сетей в образовании, это стремление ИИ достичь цели любой ценой, вплоть до «обмана».

В любом случае развитие нейронных сетей неизбежно, а разработка и применение искусственного интеллекта в России (как и в образовании) основывается, в том числе, совершенствованием правовых основ и активной работой научного и преподавательского потенциала вузов [2–5].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Мучин П.В. Мучин М.П. Экспертное тестирование возможностей нейронной сети ChatGPT-4 в русскоязычном сегменте информационного поля при решении отдельных проблем техносферной безопасности // Интерэкспо ГЕО-Сибирь. XIX Международный научный конгресс, 17–19 мая 2023 г., Новосибирск : сборник материалов в 8 т. Т. 3 : Международная научная конференция «Экономическое развитие Сибири и Дальнего Востока. Экономика природопользования, землеустройство, лесоустройство, управление недвижимостью». – Новосибирск : СГУГиТ, 2023. – С. 268-275.

2. Мучин П.В. Мучин М.П. Проблемы реализации преподавательского потенциала в аспекте цифровизации образования и внедрения онлайн-обучения // Актуальные вопросы образования. Паритет традиционного и цифрового образования в вузе: приоритеты, акценты, лучшие практики, 2–4 марта 2022 г., Новосибирск : сборник материалов Международной научно-методической конференции, Новосибирск. В 3 ч. Ч. 3. – Новосибирск : СГУГиТ, 2022. – С.7-12.

3. Перечень поручений Президента России по итогам конференции по искусственному интеллекту от 31.12.2020 г. № Пр-2242. – URL: <http://www.kremlin.ru/acts/assignments/orders/64859/>.

4. Минобрнауки России. О направлении информации. Письмо от 02.07.2021 г. № МН-5/2657. – URL: https://wuz.informio.ru/files/directory/documents/2021/07/Ps_MON_5_2657_02072021.pdf.

5. Минобрнауки России. О направлении информации. Письмо от 12.07.2021 г. № МН-5/4611. – URL: <https://docs.cntd.ru/document/608717534/>.

© П. В. Мучин, М. П. Мучин, 2025