



ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАТЕМАТИКА

В.В. Казаченок¹⁾, А.А. Русаков²⁾,

¹⁾ *Белорусский государственный университет, Беларусь*

²⁾ *Академия информатизации образования, Россия*

Аннотация. Рассматриваются возможности чат-бота ChatGPT и других больших языковых моделей (LLM) на базе искусственного интеллекта для решения математических задач различного уровня сложности. Акцентируется внимание на способностях чат-бота в обучении учащихся и студентов математике и информатике. Анализируются сильные и слабые стороны больших языковых моделей, а также риски их применения в точных науках без квалифицированного анализа ответов. Перечисляются задачи педагогов, в решении которых им может помочь рассматриваемый чат-бот, определяются направления развития LLM для повышения точности их ответов при решении математических задач.

Ключевые слова. Искусственный интеллект; чат-бот; большие языковые модели; решение задач; обучение.

ARTIFICIAL INTELLIGENCE AND MATHEMATICS

V.V. Kazachonak¹⁾, A.A. Rusakov²⁾,

¹⁾ Belarusian State University, Belarus

²⁾ Academy of Education Informatization, Russia

Abstract. The article considers the capabilities of the ChatGPT chatbot and other large language models (LLM) based on artificial intelligence for solving mathematical problems of varying complexity. The focus is on the chatbot's abilities in teaching mathematics and computer science to students. The strengths and weaknesses of large language models are analyzed, as well as the risks of their use in the exact sciences without qualified analysis of responses. The tasks of teachers that can be helped by the chatbot in question are listed, and the directions for the development of LLM are determined to improve the accuracy of their responses when solving mathematical problems.

Keywords. Artificial intelligence; chatbot; large language models; problem solving; training.

Введение

ChatGPT (от англ. Generative Pre-trained Transformer «генеративный предварительно обученный трансформер») — чат-бот с генеративным искусственным интеллектом (ИИ), разработанный компанией OpenAI и способный работать в диалоговом режиме, поддерживающий запросы на естественных языках [1]. Предусмотрено создание выходного сигнала в форме текстового ответа на заданный вопрос, который возможно преобразовать в речь, так что создается ощущение, будто вы разговариваете с кем-то в режиме реального времени.

В целом, ChatGPT и другие большие языковые модели (Large language model или LLM), доказали свою полезность для задач, отличных от генерации текста. Однако в некоторых областях их эффективность вызывает вопросы. Одной из таких областей является математика, где LLM иногда могут предоставить правильные решения для достаточно сложных задачах и, в то же время, потерпеть неудачу в тривиальных.

В частности, ChatGPT, отвечая на вопрос: насколько его возможности сравнимы с человеческими ответил: «Мои возможности в определенных областях могут быть сравнимы с возможностями человека, я не обладаю человеческими эмоциями, интуицией и способностью абстрагироваться от информации, которая может привести к ошибочным выводам. Кроме того, мои ответы могут быть ограничены тем, что я могу использовать только ту информацию, которая была доступна в моем обучающем наборе данных» [2].

ChatGPT может не обнаруживать ошибки, присущие предоставляемой ему информации. Ввиду отсутствия верификации с авторитетными источниками и какой-либо подтвержденной базой знаний модель может очень подробно и серьезно отвечать на совершенно бессмысленные вопросы, не оценивая их реалистичность [3]. Также модель может ошибаться в рассуждениях, и делать неверные выводы, хотя текст выглядит согласованным и убедительным.

Насколько можно доверять ответам ChatGPT на математические вопросы? Существенной проблемой этого чата является его уверенность в ответах, даже когда он использует ошибочную логику, утверждая, что доказательство приведено, хотя в реальности он может перепрыгивать от одного положения к другому, не выводимому на самом деле из предыдущего. Это очень затрудняет его использование в качестве надежного источника математических знаний.

Таким образом ChatGPT не разумен. Он действует только на основе наборов данных, на которых он обучен. Такие системы опираются на уже усвоенные данные и их результаты не всегда могут быть точными. Принцип их работы – не выполнение вычислений как таковых, а определение наибольшей вероятности, то есть из большого количества имеющихся данных выбираются те, чья вероятность быть точными наибольшая, но она не всегда равна 100% [4].

Основной причиной того, что чат-боты на базе ИИ испытывают трудности с математикой, является то, что они изначально не были спроектированы для этого. Учащемуся и студенту следует относиться к ответам ChatGPT не как к ответам преподавателя, а как к ответам своего одноклассника, сокурсника, понимая, что ответ может быть неправильным и необходимо перепроверять полученную информацию.

Решение задач математики и информатики

В целом, использование ChatGPT для решения математических задач открывает новые возможности и перспективы в этой области. Различные исследования показали, что ChatGPT подходит для решения математических задач школьного уровня [5].

В то же время, специалисты Беркли, оценивая поведение чат-ботов в 2023 г. в решении математических задач, указывали только на 66% правильно решенных задач. При решении задач чат-бот всегда давал какие-то ответы, и эти ответы выглядели правдоподобно. Но вы не поймете, правильный это ответ или нет, если не решите задачу сами или не знаете ответ. Пример задания чат-боту: В треугольнике ABC $AC=BC=25$, $AB=40$. Найдите $\sin A$. (Правильный ответ: 0,6).

Исследования также показали, что базовая арифметика и перестановка выражений давались боту нелегко, а если нужно было что-то доказать – он справлялся с этим. Вероятно, это связано с тем, что арифметика и использование ее правил задействуют наименьший корпус базы знаний бота, а в «доказательствах» он уже может проявить себя как языковая модель.

Что касается высшей математики, то ChatGPT обладает способностью решать математические задачи из различных областей, включая алгебраические и дифференциальные уравнения, тензорное исчисление, различные численные методы, комбинаторику, теорию вероятностей и статистику.

Для того чтобы оценивать прогресс языковых моделей, ученые используют различные эталонные тесты, или бенчмарки, которые позволяют измерить, насколько хорошо ИИ решает поставленные задачи. Среди самых популярных тестов на сегодняшний день MATH и GSM8K. На этих тестах многие передовые языковые

модели уже показывают результаты, близкие к 90% правильных ответов. При этом для чат-бота особенно трудными являются олимпиадные задачи [6].

Компания OpenAI представила новую ИИ-модель o1 для решения сложных задач. «Для сложных задач, где необходимы рассуждения, это значительный прогресс. Он представляет собой новый уровень возможностей искусственного интеллекта», — считают в компании.

По результатам тестов, o1 оказалась умнее чат-бота GPT-4о примерно в шесть раз. Она дает ответы на уровне кандидата наук при решении задач по физике, химии и биологии [7]. Что же касается возможностей ChatGPT в математике, то они остаются ниже уровня среднего аспиранта.

Для выявления по-настоящему высокого уровня понимания математики исследователям понадобился более сложный бенчмарк. Так появился FrontierMath [6]. В ходе подготовки нового теста эксперты обратились к выдающимся математикам, чтобы те предоставили свои самые трудные задачи, которые не публиковались ранее и требовали значительных усилий и глубоких знаний для их решения. Поэтому модели, рассчитывающие на успех в этом тесте, должны обладать не только большим количеством данных, но и способностью к анализу и творческому решению задач.

В результате, модели, которые демонстрировали высокие результаты на предыдущих тестах, не смогли набрать в новом тесте более 2% правильных ответов. Этот результат демонстрирует, что текущий уровень ИИ все еще далек от возможностей успешно решать самые сложные задачи, требующие значительного уровня математического мышления. Эксперты также отмечают, что появление теста FrontierMath станет важным шагом на пути к развитию ИИ, способного работать со сложными математическими задачами.

Внедрение плагинов для ChatGPT позволяет чат-боту взаимодействовать со сторонними сервисами и получать доступ к актуальной информации из интернета. Например, Wolfram. ChatGPT ранее имел возможность решать математические задачи, однако его точность и эффективность были далеко не всегда достаточными. С добавлением этого плагина, бот значительно улучшает свои возможности и точность в решении задач, используя набор вычислительных алгоритмов и базы знаний WolframAlpha.

В целом, ChatGPT способен отвечать и объяснять вопросы из широкого круга тем математики. Основная проблема всех LLM заключается в том, что они не могут исправлять ошибочные предположения и недоразумения. При более точной настройке

эти системы могут стать надежными помощниками для людей, не имеющих высшего образования в области математики. Но пока стоимость оценки результативности математических выводов LLM достаточно высока.

Какие преимущества использования моделей на базе ИИ? В первую очередь это экономия времени: задачи решаются за считанные секунды. Однако для получения наиболее точных и полных решений заданий, рекомендуется использовать ИИ в сочетании с другими инструментами и ресурсами.

В итоге, ответ на вопрос о доверии ChatGPT в точных науках не может быть однозначным. ChatGPT является мощным инструментом, способным предоставлять общую информацию, объяснения и решать типичные задачи математики. Однако, при работе со сложными и специфическими задачами, особенно требующими высокой точности, всегда рекомендуется проверять результаты.

Что касается информатики и программирования, то чат-бот может генерировать связные фрагменты кода с пояснениями для типовых задач, может находить простейшие ошибки в коде. ChatGPT также подходит для изучения языков программирования. Он грамотно пишет код, помогает разбираться в логике решения задач.

Объемную и сложную программу ChatGPT не напишет. Однако, если последовательно объяснить ему архитектуру и функциональность элементов – то он, скорее всего, поможет их достроить. Также чат-бот умеет создавать упражнения для обучения программированию с примерами решений и пояснениями.

Обучение решению задач

В настоящее время уже опубликованы статьи, посвященные использованию ChatGPT-4 в сфере образования в целом и в обучении математике в частности [8]. ChatGPT-4 обладает обширной базой знаний и демонстрирует способность доступным языком представлять темы на школьном и базовом университетском уровнях. Его возможность вести продолжительные диалоги по конкретным предметам является ценным дополнением к традиционным методам обучения, потенциально способствуя более глубокому пониманию изучаемой темы.

Но, поскольку искусственный интеллект упрощает доступ к информации, дает ответы на вопросы и демонстрирует решения задач, существует вероятность того, что учащиеся станут при любой возможности обращаться к ИИ: зачем самостоятельно искать ответы или размышлять над задачей, если это можно поручить ИИ? Как считают

исследователи, в итоге это может привести к деградации навыков критического мышления и решения проблем.

Эти опасения оправданы, так как сейчас обучающиеся получают удобный калькулятор, который, возможно, не превратит двоечника в отличника, но значительно упростит учебу и создаст атмосферу сомнений, поскольку сложно отличить ИИ-контент от результатов самостоятельной работы учащегося. Результаты исследований показали, что школьники, которые пользуются ChatGPT в обучении, хуже сдают тесты. В связи с этим доступ к ChatGPT ограничен в ряде школ России, США, Японии.

Считается, что работа, выполненная самостоятельно, помогает усваивать знания и приобретать навыки. Поэтому возникает вопрос: можно ли официально разрешить студенту использовать ИИ? Не получим ли мы специалистов с серьезными пробелами в знаниях? Опыт показывает, что запреты для студентов работают крайне плохо. Поэтому ряд ученых считают: если деятельность рутинная и может быть автоматизирована, то ее нужно отдать ИИ. Вместо того, чтобы полностью запрещать ИИ, преподаватели должны рассказывать об ответственных способах использования этой технологии.

Педагоги научили учащихся использованию математики в мире с калькуляторами. Теперь задача педагогов – научить студентов использовать новые возможности технологий. Одна из предлагаемых мер – проектная работа, причём поэтапная в течение всего курса, для оценки которой применяется формирующее оценивание. В этом случае важны несколько итераций.

Необходимо доносить до обучающихся риски и ограничения в использовании ИИ, указывать на его ошибки и слабые места. Лучше всего экспериментировать вместе с учащимися, чтобы они на собственном опыте убедились, что ИИ умеет делать, а чего пока не может.

Результаты исследований возможностей ИИ в обучении математике также показали, что при объяснении своих ответов ChatGPT часто представлял необычные или неожиданные рассуждения, даже когда приходил к правильному решению. Такой нестандартный подход может навредить студентам, изучающим стандартные математические приемы.

Исследователи считают, что образовательный потенциал ChatGPT должен быть ограничен квалифицированными математиками, которые могут проверить результаты и заметить пробелы в рассуждениях ИИ. Поэтому, менее опытные ученики должны

проявлять осторожность при самостоятельном использовании чат-бота без необходимых знаний в области математики для проверки его ответов.

Интеграция ChatGPT-4 и других LLM в образовательный процесс может привести к глубоким изменениям в подходах к преподаванию физико-математических дисциплин. Например, К.Г. Уэст утверждает, что развитие технологий LLM «требует нового аудита того, какие именно концептуальные и математические навыки должны формироваться по итогу обучения» [9; 10].

Важно трансформировать учебные планы и задания таким образом, чтобы поощрять творческое использование ИИ и его применение в роли дополнительного помощника, а не генератора рефератов и сочинений. ChatGPT хорошо объясняет базовые понятия и термины, но в то же время может что-то «придумать» или перевратить факты. Например, он может доходчиво и грамотно объяснять, как решать задачи по математике, но затрудняется применять теорию на практике. К тому же его можно запутать, даже если просто пытаться найти ошибку. Порой это приводит к парадоксальным результатам: например ChatGPT может согласиться с тем, что $2 + 2 = 5$. Так что все ответы ChatGPT нужно обязательно перепроверять.

Также ChatGPT-4 уступает в областях, требующих строгих математических доказательств и последовательных логических рассуждений. Его убедительный стиль изложения иногда может маскировать неточности, что подчеркивает необходимость использования ChatGPT-4 студентами в качестве дополнительного инструмента наряду с лекциями и учебниками, а не как единственного источника информации. В конечном итоге цель заключается не только в передаче знаний, но и в формировании навыков самостоятельного мышления и стремления к истине.

Любая языковая модель требует ведения корректного диалога. Зачастую, чтобы получить требуемое, нужно вступать в длительный и утомительный диалог. Однако ChatGPT может стать незаменимым помощником преподавателя, например, математических дисциплин в вузе, существенно облегчая рутинную техническую работу по составлению разноуровневых задач, однотипных задач в большом количестве, а также по написанию учебно-методических пособий.

Чат-бот может помочь в планировании уроков и лекций. Если, например, дать ему список источников, на основе которых нужно построить урок или целый курс, он составит структурированный план дополнив его сгенерированными практическими заданиями и тестами по теме. Также чат-бот способен оказать помощь преподавателю в

разработке учебных планов и программ для различных курсов с учетом уровня знаний студентов и требований курса.

Заключение

Сегодня ИИ упрощает процесс обучения, предоставляя быстрые и точные решения задач. Он помогает изучать различные темы, в том числе высшей математики; развивать навыки через пошаговые объяснения, что особенно важно для понимания сложных тем; подготавливать графики и визуализировать данные для лучшего понимания.

Однако ИИ не способен заменить настоящего педагога и может играть только роль вспомогательного инструмента в образовательном процессе. В настоящее время нельзя полностью полагаться на ИИ. Чат-боты – это инструмент, а не замена знаний. Рекомендуется использовать их для проверки своих решений или для изучения нового материала, но не забывать практиковаться самостоятельно.

Также исследователи отмечают, что поведение чат-ботов может существенно изменяться за относительно короткий промежуток времени, поэтому необходим постоянный контроль за качеством работы алгоритмов искусственного интеллекта.

Ключевым направлением развития LLM является уменьшение ошибок и случайных «галлюцинаций» в ответах. Чтобы добиться этого, модели начнут чаще обращаться к различным внешним источникам информации.

Список библиографических ссылок (на языке оригинала)

1. Нейросеть Chat GPT на русском. GPT-chatbot [Электронный ресурс]. URL: <https://gpt-chatbot.ru/> (дата обращения: 23.05.2025).
2. Логачева О.М., Логачев А.В. ChatGPT как цифровой помощник в обучении математическим дисциплинам [Электронный ресурс]. URL: <https://sgugit.ru/upload/science-and-innovations/conference-ssga/international-scientific-methodical-conference-actual-problems-of-education/collections-of-materials-nmk-2023/part1/088-094.pdf> (дата обращения: 23.05.2025).
3. Кузнецов А. Обзор ChatGPT с примерами использования в различных задачах. *Искусственный интеллект; исследования и прогнозы в IT* [Электронный ресурс]. URL: <https://habr.com/ru/articles/708646/> (дата обращения: 23.05.2025).
4. Искусственный интеллект умеет писать стихи, но все еще испытывает трудности с математикой. *Онлайн-школа программирования для детей* [Электронный

ресурс]. URL: <https://code-it-school.ru/blog/iskusstvennyj-intellekt-umeet-pisat-stihi-no-vse-eshhe-ispytyvaet-trudnosti-s-matematikoj/> (дата обращения: 23.05.2025).

5. Проверяем, какой ИИ бот лучше решает математические задачи; сравниваем ChatGPT и Gemini на задачах уровня ЕГЭ. *Математика; машинное обучение* [Электронный ресурс]. URL: <https://habr.com/ru/sandbox/220498/> (дата обращения: 23.05.2025).

6. Левченко С. ИИ все еще слаб в высшей математике: неутешительные выводы ученых. hi-tech 2024 [Электронный ресурс]. URL: <https://hi-tech.mail.ru/news/117644-ii-vse-eshe-slab-v-vysshej-matematike-neuteshitelnye-vyvody-uchenyh/> (дата обращения: 23.05.2025).

7. Пантох С. OpenAI представила новую ИИ-модель o1 для решения сложных задач. hi-tech 2024 [Электронный ресурс]. URL: <https://incrypted.com/openai-predstavila-novuyu-ii-model-o1-dlya-resheniya-slozhnyh-zadach/> (дата обращения: 23.05.2025).

8. Мариносян А.Х. ChatGPT-4 в обучении физике и математике: возможности, ограничения и перспективы совершенствования. *Вестник МГПУ. Серия «Информатика и информатизация образования»*. 2024; 4(70):95–115.

9. Казаченок В.В., Русаков А.А. Искусственный интеллект в электронном обучении. Электронный науч.-методич. журнал «Педагогика информатики». 2024; 1-2. URL: https://pcs.bsu.by/2024_1-2/1ru.pdf (дата обращения: 23.05.2025).

10. West C.G. Advances in apparent conceptual physics reasoning in ChatGPT-4 [Electronic resource]. URL: <https://arxiv.org/abs/2303.17012>. (Date of access: 23.05.2025).

References (на английском языке)

1. Neyroset' Chat GPT na russkom. GPT-chatbot [Electronic resource]. Available at: <https://gpt-chatbot.ru/>. (Date of access: 23.05.2025). (In Russian)
2. Logacheva O.M., Logachev A.V. ChatGPT kak tsifrovoy pomoshchnik v obuchenii matematicheskim distsiplinam [Electronic resource]. Available at: <https://sgugit.ru/upload/science-and-innovations/conference-ssga/international-scientific-methodical-conference-actual-problems-of-education/collections-of-materials-nmk-2023/part1/088-094.pdf>. (Date of access: 23.05.2025). (In Russian)
3. Kuznetsov A. Obzor ChatGPT s primerami ispol'zovaniya v razlichnykh zadachakh. *Iskusstvennyj intellekt; issledovaniya i prognozy v IT* [Electronic resource]. Available at: <https://habr.com/ru/articles/708646/>. (Date of access: 23.05.2025). (In Russian)

4. Iskusstvennyy intellekt umeyet pisat' stikhi, no vse yeshche ispytyvayet trudnosti s matematikoy. *Onlayn-shkola programmirovaniya dlya detey* [Electronic resource]. Available at: <https://code-it-school.ru/blog/iskusstvennyj-intellekt-umeet-pisat-stihi-no-vse-eshhe-ispytyvaet-trudnosti-s-matematikoj/>. (Date of access: 23.05.2025). (In Russian)
5. Proveryayem, kakoy II bot luchshe reshayet matematicheskiye zadachi; sravnivayem ChatGPT i Gemini na zadachakh urovnya YEGE. *Matematika; mashinnoye obuchenije* [Electronic resource]. Available at: <https://habr.com/ru/sandbox/220498/>. (Date of access: 23.05.2025). (In Russian)
6. Levchenko S. II vse yeshche slab v vysshey matematike: neuteshitel'nyye vyvody uchenykh. hi-tech 2024 [Electronic resource]. Available at: <https://hi-tech.mail.ru/news/117644-ii-vse-eshe-slab-v-vysshej-matematike-neuteshitelnye-vyvody-uchenyh/>. (Date of access: 23.05.2025). (In Russian)
7. Pantyukh S. OpenAI predstavila novyyu II-model' o1 dlya resheniya slozhnykh zadach. hi-tech 2024 [Electronic resource]. Available at: <https://incrypted.com/openai-predstavila-novyyu-ii-model-o1-dlya-resheniya-slozhnyh-zadach/>. (Date of access: 23.05.2025). (In Russian)
8. Marinosyan A.KH. ChatGPT-4 v obuchenii fizike i matematike: vozmozhnosti, ograniceniya i perspektivy sovershenstvovaniya [ChatGPT-4 in teaching physics and mathematics: capabilities, limitations, and prospects for improvement]. *Vestnik MGPU. Seriya «Informatika i informatizatsiya obrazovaniya»*. 2024; 4(70):95– 115. (In Russian)
9. Kazachenok V.V., Rusakov A.A. Iskusstvennyy intellekt v elektronnom obuchenii. *Elektronnyy nauch.-metodich. zhurnal «Pedagogika informatiki»*. 2024; 1-2. Available at: https://pcs.bsu.by/2024_1-2/1ru.pdf. (Date of access: 23.05.2025). (In Russian)
10. West C.G. Advances in apparent conceptual physics reasoning in ChatGPT-4 [Electronic resource]. Available at: <https://arxiv.org/abs/2303.17012>. (Date of access: 23.05.2025).