

УДК 004.853

***ПРИМЕНЕНИЕ ЧАТ-БОТОВ НА ОСНОВЕ БОЛЬШИХ ЯЗЫКОВЫХ
МОДЕЛЕЙ ДЛЯ ПРЕПОДАВАНИЯ ДИСЦИПЛИНЫ «НЕЙРОСЕТЕВЫЕ
ТЕХНОЛОГИИ АНАЛИЗА ДАННЫХ» В ВЫСШЕЙ ШКОЛЕ***

Рожков В.А.

аспирант,

Забайкальский государственный университет,

Чита, Россия

Аннотация

Обсуждается применение чат-ботов на основе больших языковых моделей (таких, как ChatGPT и подобных) в контексте преподавания дисциплины «Нейросетевые технологии анализа данных» в высшей школе. Для этого рассматриваются существующие проблемы при преподавании данной дисциплины. Она является продолжением курса по машинному обучению и включает в себя исследование и обработку данных для их использования в искусственных нейронных сетях (ИНС). Сделаны заключения по тому, способен ли чат-бот решить эти проблемы и каким образом он может повлиять на образовательный процесс по данной дисциплине. Даны рекомендации по улучшению преподавания нейросетей в условиях развития и распространения интеллектуальных чат-ботов.

Ключевые слова: Преподавание, чат-бот, большие языковые модели, машинное обучение, искусственные нейронные сети.

***THE USE OF LARGE LANGUAGE MODEL-BASED CHATBOTS FOR
TEACHING THE DISCIPLINE "NEURAL NETWORK TECHNOLOGIES OF
DATA ANALYSIS" IN HIGHER EDUCATION***

Rozhkov V.A.

postgraduate student,

Transbaikal State University,

Chita, Russia

Abstract

The use of chatbots based on large language models (such as ChatGPT and the like) in the context of teaching the discipline "Neural network technologies of data analysis" in higher education is discussed. The existing problems in teaching this discipline are considered. This discipline is a continuation of the machine learning course and includes the exploratory analysis and processing of data to use it in artificial neural networks (ANN). Conclusions are drawn on whether a chatbot is capable of solving these problems and how it can influence the educational process in this discipline. Recommendations are given on improving the teaching of neural networks in the context of the development and spread of intelligent chatbots.

Keywords: Teaching, chatbot, large language models, machine learning, artificial neural networks.

Актуальность темы заключается в широком распространении чат-ботов на основе больших языковых моделей и их адаптации студентами в учебной деятельности, а также в необходимости сделать такой многогранный и сложный предмет, как «Нейросетевые технологии анализа данных» более доступным для преподавания и формирования компетенций по нему.

Данная дисциплина направлена на формирование следующих компетенций [10]:

– Понимание математического аппарата, на основе которого строятся нейронные сети, понимание их фундаментальных элементов и алгоритмов, умение работать с нейросетями различных архитектур;

– Умение обрабатывать данные и создавать на их основе полезные признаки для обучения нейросетей, обучать и оценивать их точность, осуществлять их тонкую настройку для получения высокой эффективности, изучать влияние входных признаков на работу модели и ее результаты;

– Умение применять нейросети для решения прикладных задач, развертывать их в рабочей среде, осуществлять их мониторинг;

– Знание передовых разработок в области вычислительного интеллекта.

Данная инженерная дисциплина была сформирована в связи с потребностью современной экономики в специалистах, умеющих работать с большими данными и моделями вычислительного интеллекта, и готовых к эффективной работе в современном информационном мире [2]. При этом инженерная практика оказывает прямое и важное влияние на инженерную педагогику, что обуславливает ряд особенностей при преподавании нейросетей [3].

Несмотря на то, что роль искусственного интеллекта в образовании уже изучалась, существует недостаток исследований о том, как чат-боты на базе больших языковых моделей влияют на процесс обучения по машинному обучению, нейронным сетям и технологиям анализа и обработки данных, т.е. на те разделы, которые присутствуют в дисциплине «Нейросетевых технологий анализа данных».

Основными вопросами данного исследования являются следующие:

1. Как устроено обучение нейросетевым технологиям анализа данных сейчас и какие есть проблемы у этого процесса?

2. Каким образом внедрение чат-бота способно решить эти проблемы?

Ответы на эти вопросы позволят оценить достоинства и недостатки активного внедрения чат-ботов для решения проблем педагогического процесса в данной дисциплине.

Интеллектуальный помощник, работающий на основе большой языковой модели (англ. «large language model», LLM), или чат-бот, является полезным инструментом в учебной и профессиональной деятельности программистов и специалистов по исследованию данных. Чат-бот способен в произвольной форме дать пояснение сложного термина, привести пример кода, найти ошибки в уже написанном коде, написать для него документацию или даже предоставить интеграционные и юнит-тесты для обеспечения покрытия кодовой базы тестами.

Чат-бот является диалоговым инструментом, поскольку он позволяет осуществлять информативное взаимодействие между коммуницирующими сторонами (в данном случае – между человеком и машиной), посредством которого возникает понимание [1]. Со стороны машины понимание является результатом обучения нейросети с подкреплением (от англ. reinforcement learning) – подобный процесс обучения близок к тому, как обучаются люди. Искусственные нейроны по отдельности не обладают способностями к человекоподобному диалогу, но при достаточно большом количестве нейронов у сети появляется свойство эмерджентности [9], что в итоге приводит к созданию естественной обратной связи и приращению информации между пользователем и нейросетью.

Современные чат-боты достигли уровня, на котором они могут адаптировать свои ответы к различным контекстам, например, к разным предметным областям, и по желанию делать объяснения простыми или более развернутыми, а также способны совместно с человеком генерировать идеи, решать проблемы и исследовать темы, давая соответствующие ответы.

Важно отметить, что чат-бот не способен полностью заменить человека и его навыки. Он может только помочь ему с решением конкретных, четко поставленных вопросов, т.к. на текущем этапе развития исследовательские навыки LLM оставляют желать лучшего ввиду следующих причин:

– Отсутствие способности к творческому мышлению;

– Отсутствие детерминированности вывода, т.е. при отправлении чат-боту одного и того же текстового запроса (или промпта, от англ. «prompt») он может возвращать разные результаты, при этом не давая пояснений, чем обусловлена эта разница в ответах;

– Выдаваемые большой языковой моделью данные и факты могут частично или полностью вымышленными, т.е. в своих ответах чат-бот может выдавать ложные факты за истинные, особенно при увеличении объема промпта. Такое явление называют «галлюцинациями» или «деменцией» нейросети.

Из этого следует, что ответственность за код и алгоритм решения задачи всегда лежит на плечах специалиста, использующего сгенерированный код, и любая информация, полученная от чат-бота, должна быть перепроверена человеком.

Потребность в использовании чат-ботов студентами, обучающимся по дисциплине «Нейросетевые технологии анализа данных» обусловлена не только желанием ускорить и автоматизировать разработку программного обеспечения и нейросетей, но и потребностью осваивать новые, насыщенные сложными концепциями темы в условиях ограниченного числа учебных занятий и стремительного развития отрасли вычислительного интеллекта.

Внедрение чат-бота в работу преподавателя соответствует необходимости поддерживать учебную программу в соответствии с вектором развития области и наиболее востребованными разработками в быстро развивающейся науке о данных.

Особенности подготовки специалистов по исследованию данных заключаются в следующем:

– Высокий порог входа в предмет. Помимо понимания базовых моделей машинного обучения, требуется понимание математических операций, лежащих в основе искусственных нейронных сетей, знание статистики,

владение навыками программирования и умение работать со структурами и алгоритмами обработки данных;

- Сложность и многообразие науки о данных (англ. «data science», DS);
- Необходимости качественной практики через выполнение объемных заданий в течение ограниченного времени учебного курса.

Далее выделены основные сложности, возникающие при обучении и при профессиональной работе с искусственными нейронными сетями и технологиями анализа данных [2, 8, 12]:

1. Сфера вычислительного интеллекта в настоящее время проходит через стадию бурного роста, и из-за этого появляется много разных информационных ресурсов для изучения, разной степени качества и сложности. Содержание обучающих курсов по нейросетям постоянно расширяется новыми сложными темами, при этом объем информации, требуемой для усвоения только увеличивается. Из-за этого появляется сложность в обучении в виде большого количества неактуализированной, неконцептуализированной информации, т.е. учебная информация не в полной мере трансформируется в знания и компетенции;

2. Для эффективной работы нейросетям требуются правильно обработанные данные, т.е. существует сложная задача в виде получения и обработки репрезентативных данных из предметной области для модели. Этап предварительного анализа и обработки данных необходим, т.к. в «сырых», исходных данных зачастую могут скрываться шумы и/или ошибки (в т.ч. смысловые) вследствие помех или неправильной записи. При этом одни и те же данные могут быть представлены разными способами, и только эмпирическим путем можно выяснить, какой способ представления лучше всего способствует достаточно точной работе модели;

3. Бывают ситуации, когда набор данных не содержит достаточное количество информации для возможности качественного обобщения. В этом случае приходится прибегать к аугментации данных, т.е. к генерации новых

данных на основе старых с применением известных преобразований (к примеру, для растровых изображений это поворот, увеличение, сжатие и т.д.), либо к созданию синтетических данных, имитирующих наблюдения из реального мира. При этом нужно учитывать, что подобные манипуляции могут снизить общее качество данных и повысить риск переобучения модели, особенно если они недостоверно отражают предметную область;

4. Сложности, связанные с переобучением и недообучением модели, и нахождением баланса между этими двумя крайностями для получения адекватной предсказательной силы (обобщающей способности) модели. Важно, чтобы модель была не слишком простой (неспособной отразить закономерности в данных), но и не слишком сложной (не была чересчур чувствительна к колебаниям в данных), и для этого требуется правильная настройка механизмов регуляризации внутри модели;

5. Алгоритмы вычислительного интеллекта, а в особенности нейросети, сами по себе являются сложными не только с математической точки зрения, но и в плане понимания исследователем того, каким образом исходные данные трансформируются в прогноз посредством модели, и определения, насколько этот прогноз достоверен, т.е. часто возникают сложности с пониманием математической сущности алгоритма и его возможных результатов ввиду недостатка качественной практики;

6. Бывает тяжело соотнести эффективность конкретного алгоритма с определенным набором данных, т.е. нет конкретной формулы, способной ответить на вопрос, можно ли добиться высокой эффективности от определенной комбинации из модели и набора данных за разумное время и усилия. Есть только эвристический алгоритм для перебора признаков, способных повлиять на результаты модели. Понимание ограничений, допущений модели и набора данных имеет решающее значение при определении, соотносятся ли предсказания модели с наблюдениями из реального мира;

7. Сложность с интерпретированием работы моделей, из которой вытекает проблема доверия к моделям и их результатам. Как правило, прогностическая способность модели обратно пропорциональна прозрачности ее работы, и это в первую очередь относится к искусственным нейронным сетям – эти модели являются непрозрачными. Существует ряд алгоритмов, способных объяснять влияние признаков как на работу модели, так и на ее конечный результат. Тем не менее, пока что среди разнообразных методов объяснения моделей вычислительного интеллекта не выделен единый стандарт, и каждый из этих методов имеет свои ограничения, зависящие в том числе от предметной области и способа представления ее данных в модели;

8. Сложный цикл отладки моделей вычислительного интеллекта, который на порядок сложнее, чем в классических программах. Ввиду отсутствия сложившегося инструментария для отладки и тестирования нейросетей их разработчикам очень легко допустить ошибку;

9. Необходимость доступа к высокопроизводительному оборудованию, т.е. эффективное обучение модели осложняется высокими требованиями к вычислительным ресурсам. Чем крупнее модель и чем больше набор данных, тем больше памяти, вычислений и энергии требуется для обучения модели и ее тонкой настройки, подбора гиперпараметров;

10. Тонкая настройка гиперпараметров модели требует от разработчика нейросетей особый вид интуиции. При работе с искусственными нейронными сетями нет строгого способа заранее узнать, какая именно архитектура сети (тип, количество слоев и нейронов в них, функции активации слоев, функция потерь, алгоритм оптимизации при обучении и др.) будет наиболее эффективна для решения поставленной задачи, поэтому требуется ряд вычислительных экспериментов для установления подходящей архитектуры;

11. Даже при решении всех вышеперечисленных сложностей в конце концов остается проблема внедрения разработанной для конкретной предметной области нейросети в качестве помощника для человека-Дневник науки | www.dnevnika.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

специалиста. Она обусловлена тем, что модель предоставляет прогноз, который точен с определенной вероятностью, и не предоставляет развернутую аргументацию, нарратив или мотивацию, понятную человеку, за каждым прогнозом. В таком случае прогноз может быть проигнорирован, даже если он окажется верным. Подобная ситуация называется проблемой Кассандры [11]. Следовательно, специалисту, работающему в своей области требуется хорошее понимание принципов работы ИИ для его эффективного использования.

Эти сложности отражены в виде диаграммы на рисунке 1.

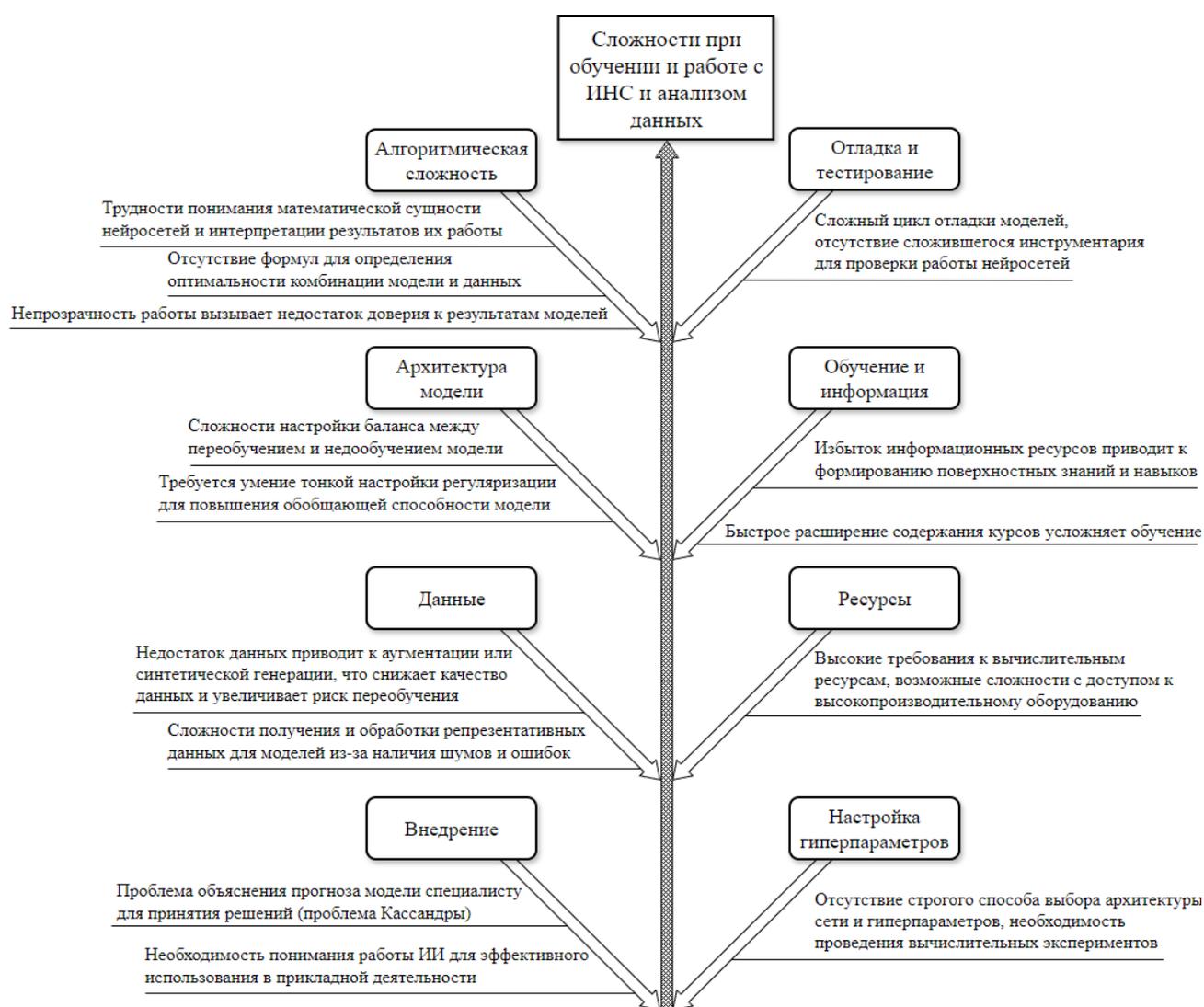


Рис. 1 – Сложности при обучении и работе с ИИС и анализом данных (рисунок составлен автором)

В статье [6] проведен SWOT-анализ использования чат-бота (а именно – ChatGPT) в образовательном процессе. Рассмотрим специфику чат-ботов применительно к дисциплине «Нейросетевые технологии анализа данных»:

– На данный момент чат-боты не могут самостоятельно подобрать наиболее эффективную архитектуру нейросети для заданного набора данных, они только лишь могут указать на возможные способы улучшения эффективности модели, а также привести примеры использования тех или иных моделей или вспомогательных методов;

– Весь код, полученный от чат-бота должен быть обязательно перепроверен человеком на синтаксические и смысловые ошибки. При усложнении условий задачи вероятность получения ошибок возрастает;

– Для преподавателя чат-бот может послужить способом упростить такие рутинные моменты, как разработка индивидуализированных заданий, обновление или подготовка новых примеров кода для практических занятий, или автоматизация написания детальной обратной связи по выполненным заданиям [7]. Также он может оказывать помощь при подготовке лекционных занятий;

– Преподавателю при принятии работ приходится задавать больше вопросов по теории и чаще уточнять детали программной реализации, но при этом времени для занятий не становится больше. Но при этом нельзя делегировать проверку работ чат-боту полностью – машина никогда не сможет заменить живую учебную коммуникацию;

– При должном использовании студентами чат-бот может давать им рекомендации по решению задачи, искать ошибки в коде и исправлять их, предоставлять ответы на вопросы, не требующие вмешательства преподавателя, позволяя не дожидаться занятия и продолжать самостоятельную работу [7]. Это соответствует потребности в преодолении разрыва между знаниями и их практическим применением, что является важнейшим условием подготовки

конкурентоспособного инженера [4], т.е. чат-бот может способствовать развитию самостоятельности;

– Внедрение чат-бота в образовательный процесс, с одной стороны, повышает доступность информации, давая возможность не дожидаться встречи с преподавателем, чтобы задать вопросы, но с другой стороны полученная от чат-бота информация необязательно трансформируется в знания и навыки. При этом сгенерированный ответ может содержать смысловые ошибки или неточности, т.е. будучи на первый взгляд адекватным, ответ может быть в корне неверным, т.к. принцип работы чат-ботов заключается в подборе наиболее вероятного ответа, что ставит их ответы на ступень ниже по сравнению с ответами поисковых систем, которые оснащены технологией «knowledge graph» (семантической базой данных). Кроме того, не все чат-боты способны давать ссылки на источники информации, не говоря уже об эффективном поиске релевантной информации, что дополнительно снижает уровень доверия к их ответам. Несмотря на то, что чат-бот способен «разжевать» отдельные концепции и помочь с заданиями, он подвергает своего пользователя большому риску академической непорядочности [5], т.к. высокое качество ответов чат-бота приводит к переоценке его возможностей, к полному делегированию учебных обязанностей или некорректному использованию сгенерированного ответа [6];

– У чат-бота могут отсутствовать информация и примеры кода по новейшим (или редко используемым) разработкам в области искусственного интеллекта, т.к. его база данных, в отличие от поисковых систем, обновляется с некоторой задержкой, и может не учитывать все доступные источники информации. А при анализе файлов, содержащих научные тексты, может выдавать предвзятую информацию вместо научных фактов из-за сложного научного языка, недостающих базовых знаний и неумения интерпретировать результаты научных работ.

Независимо от отношения к интеллектуальным текстовым помощникам, стоит признать, что они оказывают существенное влияние на качество образования и нужно начать прикладывать усилия для того, чтобы этот инструмент не скомпрометировал формирование компетенций у обучающихся, лишив их самостоятельности мышления.

При благоразумном использовании чат-бот имеет потенциал для помощи в изучении темы или в осознанном выполнении работы студентами. Чат-бот является инструментом для самостоятельного изучения тем, изложенных в курсе дисциплины «Нейросетевых технологий анализа данных».

С другой стороны, чат-бот – это еще один способ выполнить задание «для галочки», если студент во многом или полностью полагается на сгенерированные ответы и код от интеллектуального помощника. Из-за этого преподавателю требуется задавать больше вопросов по заданиям, чтобы убедиться, делал ли работу сам студент или нет. При этом усложнение условий преподавания никак не отражается на учебном плане – количество занятий остается таким же, как и до повсеместного появления чат-ботов. Для решения этой проблемы можно принять меры по увеличению числа занятий, либо по введению элективных курсов или факультативных занятий, а также следует расширить преподавательский состав специалистами по вычислительному интеллекту [2].

Основной проблемой в преподавании нейросетевых технологий является большое разнообразие материалов курса и необходимость решения объемных заданий при ограниченном числе занятий. Чат-бот может выступить как инструмент для частичного решения этой проблемы в качестве самопомощи студентам, позволяя прорабатывать объемную теоретическую информацию. При этом чат-бот несет в себе угрозу для формирования знаний, навыков и самостоятельности студентов в виде академической неупорядоченности.

Пока что тяжело дать ясный ответ на вопрос, действительно студенты чему-то учатся при использовании чат-ботов, а не просто слепо полагаются на

Дневник науки | www.dnevnika.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

сгенерированные ответы, т.к. их распространенность неизбежно увеличивается и студенты все чаще используют их в своей деятельности, впрочем, как и специалисты, профессионально работающие в области. Важно иметь в виду, что любая информация, полученная от чат-бота, должна быть обязательно перепроверена.

Библиографический список

1. Игумнова Е.А. Диалогическое взаимодействие – вектор развития современного образования // Педагогический ИМИДЖ. 2016. № 3(32). С. 8-15.

2. Каспаров И.В. Проблемы и перспективы изучения дисциплины «Системы искусственного интеллекта» на примере филиала СамГУПС в г. Нижнем Новгороде // Бизнес. Образование. Право. 2024. № 2(67). С. 486-491. DOI: 10.25683/VOLBI.2024.67.994

3. Мартынов В.Г., Шейнбаум В.С. Инженерная педагогика в контексте инженерной деятельности // Высшее образование в России. 2022. Т. 31. № 6. С. 152-168. DOI: 10.31992/0869-3617-2022-31-6-152-168

4. Мустафина Д.А., Рахманкулова Г.А., Короткова Н.Н. Модель конкурентоспособности будущего инженера-программиста // Современные наукоемкие технологии. 2010. № 8. С. 16-20.

5. Середкина Е.В., Мезин Е.В. Как может повлиять ChatGPT на культуру диалога и образование? // Научно-исследовательские исследования. 2023. № 3. С. 74-89. DOI: 10.31249/scis/2023.03.04

6. Токтарова В.И., Ребко О.В. ChatGPT в работе педагога: возможности и риски использования // Цифровая гуманитаристика и технологии в образовании (DHTE 2023): сб. статей IV Международной научно-практической конференции. 16–17 ноября 2023 г. / Под ред. В.В. Рубцова, М.Г. Сороковой, Н.П. Радчиковой. М.: Издательство ФГБОУ ВО МГППУ, 2023. С. 421–430.

7. Широколобова А.Г. Искусственный интеллект как инструмент оптимизации работы преподавателя высшей школы // Педагогика. Вопросы теории и практики. 2024. Т. 2. № 9. С. 138-145.

8. Barbierato E., Gatti A. The Challenges of Machine Learning: A Critical Review // Electronics. 2024. Vol. 13. № 416. DOI: 10.3390/electronics13020416

9. Хабр – Нейронные сети, графы и эмерджентность [электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/751340> (дата обращения: 22.12.24).

10. ivtipm/ML – Нейросетевые технологии анализа данных [электронный ресурс]. – Режим доступа: https://github.com/ivtipm/ML/blob/main/plans/2024_NN.md (дата обращения: 23.12.24).

11. The Science Times – AI in Medicine Encounters Cassandra Complex; Do We Ignore Accurate Predictions From Machine Learning Algorithms? [электронный ресурс]. – Режим доступа: <https://www.sciencetimes.com/articles/46640/20231020/ai-medicine-encounters-cassandra-complex-ignore-accurate-predictions-machine-learning.htm> (дата обращения: 25.12.24).

12. Zayd's Blog – Why is machine learning 'hard'? [электронный ресурс]. – Режим доступа: <https://ai.stanford.edu/~zayd/why-is-machine-learning-hard.html> (дата обращения: 01.12.24).

Оригинальность 75%