

УДК 004.8

Бегдесенова Диана Амиркызы, ученица Республиканской Физико-
Математической школы, г. Алматы

СРАВНИТЕЛЬНЫЙ АНАЛИЗ НЕЙРОСЕТЕЙ И ИХ ЭФФЕКТИВНОСТИ В РАЗЛИЧНЫХ ЗАДАЧАХ

Аннотация. В статье представлено исследование современных нейросетевых систем и анализ их эффективности при решении различных учебных и логических задач. Актуальность работы обусловлена активным внедрением нейросетей в образовательную сферу и необходимостью осознанного выбора инструментов искусственного интеллекта для учебной деятельности учащихся. Цель исследования заключается в сравнительном анализе возможностей популярных нейросетей и определении их сильных и слабых сторон. В ходе работы были использованы методы сравнительного анализа, экспериментального тестирования и балльной оценки результатов. Нейросети оценивались по способности разгадывать шифры, анализировать тексты, писать эссе, решать математические задачи, отвечать на исторические вопросы и генерировать программный код. По итогам исследования составлен рейтинг нейросетей и сделан вывод о том, что универсальной нейросети, одинаково эффективно выполняющей все типы заданий, не существует, а каждая модель имеет собственную область наибольшей эффективности.

The article presents a study of modern neural network systems and an analysis of their effectiveness in solving various educational and logical tasks. The relevance of the research is determined by the rapid integration of neural networks into the educational sphere and the need for a conscious selection of artificial intelligence tools for students' learning activities. The purpose of the study is to conduct a comparative analysis of popular neural networks and to identify their strengths and weaknesses. The research methods include comparative analysis, experimental testing, and a score-based evaluation system. Neural networks were assessed based on their ability to decode ciphers, analyze texts, write essays, solve mathematical

problems, answer historical questions, and generate program code. As a result, a ranking of neural networks was compiled, and it was concluded that there is no universal neural network equally effective for all types of tasks.

Ключевые слова: нейросети, искусственный интеллект, образовательные технологии, сравнительный анализ, машинное обучение, генеративные модели

Keywords: neural networks, artificial intelligence, educational technologies, comparative analysis, machine learning, generative models

Введение

В последние годы технологии ИИ и нейросетей получили широкое распространение и стали активно использоваться в различных сферах жизни, в том числе и в образовании. Современные нейросети способны анализировать тексты, решать задачи, писать программный код и отвечать на вопросы по разным учебным предметам. В связи с этим многие школьники начинают использовать нейросети как инструмент для обучения и подготовки к урокам. Однако при активном использовании нейросетей возникает проблема достоверности полученных результатов. Несмотря на внешнюю убедительность ответов, нейросети могут допускать логические, фактические и вычислительные ошибки. Особенно это проявляется при решении математических задач, анализе исторических событий и выполнении сложных логических заданий. В связи с большим количеством существующих нейросетей становится актуальным вопрос: **какие из них наиболее эффективны для выполнения учебных задач и можно ли считать их равноценными.**

Актуальность данной работы заключается в необходимости объективного сравнительного анализа различных нейросетей, используемых учащимися в

учебной деятельности. Результаты исследования могут помочь школьникам осознанно и безопасно использовать нейросети, понимая их возможности и ограничения.

Цель исследования - сравнить эффективность работы различных нейросетей при выполнении учебных и интеллектуальных заданий разного типа.

Для достижения цели были поставлены следующие **задачи**:

1. Изучить понятие нейросетей и принципы их работы.
2. Проанализировать возможности применения нейросетей в образовательной среде.
3. Подобрать задания разных типов для практического исследования.
4. Провести эксперимент с использованием нескольких нейросетей.
5. Проанализировать полученные ответы и выявить типичные ошибки.
6. Сделать выводы о сильных и слабых сторонах каждой нейросети.

Объект исследования - нейросети и системы искусственного интеллекта.

Предмет исследования - качество и корректность ответов нейросетей при выполнении учебных заданий.

Гипотеза исследования заключается в том, что разные нейросети демонстрируют различный уровень эффективности в зависимости от типа задания, и не существует универсальной нейросети, одинаково хорошо справляющейся со всеми видами учебных задач.

В ходе работы использовались следующие **методы исследования**: анализ, сравнение, эксперимент, обобщение и систематизация результатов.

Практическая значимость работы состоит в возможности использования полученных результатов школьниками и учителями для более осознанного применения нейросетей в учебном процессе.

Глава 1. Теоретические основы нейросетей

1.1. Понятие и принципы работы нейросетей

Нейросеть - это модель искусственного интеллекта, работа которой основана на принципах функционирования человеческого мозга. Она состоит из множества взаимосвязанных элементов - искусственных нейронов, которые обрабатывают информацию и передают её дальше по сети. Нейросети обучаются на больших объёмах данных, выявляя закономерности и связи между элементами информации.

Современные языковые модели нейросетей обучаются на текстах из различных источников и способны генерировать связные и логически оформленные ответы. Однако нейросеть не «понимает» информацию в человеческом смысле, а лишь прогнозирует наиболее вероятное продолжение текста на основе предыдущего контекста.

1.2. Использование нейросетей в образовании

В образовательной сфере нейросети используются для объяснения учебного материала, подготовки к контрольным работам, написания сочинений и эссе, решения различных задач и программирования. К основным преимуществам нейросетей относятся высокая скорость ответа, доступность и возможность получения объяснений в удобной форме.

В то же время использование нейросетей несёт определённые риски. Нейросети могут предоставлять устаревшую или некорректную информацию, допускать логические ошибки и уверенно выдавать неверные ответы за правильные. Это особенно опасно в учебной среде, где учащиеся могут принять ошибочный ответ за верный.

1.3. Ограничения и причины ошибок нейросетей

Одной из основных проблем нейросетей является так называемый эффект «галлюцинаций», при котором модель генерирует правдоподобный, но фактически неверный ответ. Кроме того, нейросети испытывают трудности при решении математических задач, так как такие задания требуют строгого логического вывода и точных вычислений, а не вероятностного предсказания.

Также на качество ответа сильно влияет формулировка запроса. Недостаточно точные или двусмысленные запросы могут приводить к ошибочным результатам, что делает необходимым умение правильно формулировать вопросы при работе с нейросетями.

Глава 2. Методика и организация исследования

2.1. Выбор нейросетей для исследования

Для проведения исследования были выбраны наиболее популярные и доступные нейросети и ИИ-ассистенты, используемые учащимися: ChatGPT, Gemini, DeepSeek, Алиса AI, Claude, Perplexity AI, Microsoft Copilot, Notion AI, GigaChat и MiMo-V2. Такое количество нейросетей позволило получить более объективные и наглядные результаты сравнения.

2.2. Условия проведения эксперимента

Эксперимент проводился в одинаковых условиях для всех нейросетей. Каждой модели предлагались идентичные запросы без изменения формулировок. В ряде заданий допускалось использование подсказок, при этом фиксировалось количество попыток и время ответа. Все нейросети использовались в бесплатной версии, что также влияло на результаты исследования.

2.3. Критерии оценки результатов

Оценка ответов нейросетей проводилась по следующим критериям:

- правильность и точность ответа;
- логичность рассуждений;
- полнота раскрытия темы;
- присутствие или отсутствие ошибок;
- понятность и структурированность;
- удобство использования для учащихся.

Полученные данные фиксировались и анализировались, после чего были сделаны выводы о степени эффективности каждой нейросети.

2.4. Балльная система оценки результатов

Для объективного сравнения эффективности нейросетей в рамках данного исследования была введена **единая 10-балльная система оценки**, позволяющая количественно оценить качество выполнения каждого задания.

Каждое задание оценивалось по следующим критериям:

1. **Правильность ответа** - 4 балла
2. **Логика рассуждений и объяснение решения** - 2 балла
3. **Полнота и детализация ответа** - 2 балла
4. **Ясность и структурированность** - 1 балл
5. **Время получения корректного ответа** - 1 балл

3.1. Анализ способности нейросетей к разгадыванию шифров

Целью данного этапа исследования было определить, насколько эффективно различные нейросети справляются с задачами на логический анализ и распознавание шифров. Для этого были использованы задания, связанные с **шифром Цезаря**, включая как русскоязычные, так и англоязычные тексты.

Нейросетям предлагались следующие задания:

- определить тип шифра;
- установить величину сдвига;
- расшифровать зашифрованный текст.

В качестве исходных данных использовались как открытые предложения, так и зашифрованные тексты, полученные путём сдвига алфавита вправо на 5 букв. В ходе эксперимента фиксировались:

- количество попыток;
- необходимость использования подсказок;
- время, затраченное на получение правильного ответа;

- корректность расшифровки.

Ход и результаты эксперимента

В результате исследования были получены следующие данные.

ChatGPT - не смог корректно определить шифр с первой попытки и изначально лишь предположил возможный вариант. После подсказки о том, что используется шифр Цезаря, нейросеть смогла определить тип шифра, однако ошиблась в выборе сдвига. Только с третьей попытки ChatGPT правильно установил тип шифра и величину сдвига. Общее время рассуждений составило около 26 секунд на первую попытку и около 20 секунд после получения подсказки.

Gemini (быстрая модель) - потребовала около пяти попыток для подтверждения использования шифра Цезаря, при этом точный сдвиг был определён не сразу.

Gemini (думающая модель) - показала лучший результат: спустя 54 секунды нейросеть правильно определила тип шифра и величину сдвига, продемонстрировав более последовательное рассуждение.

DeepSeek - в течение более двух минут не смог определить правильный тип шифра и сдвиг. Даже после двух подсказок (о типе шифра и величине сдвига) нейросеть не дала корректного ответа.

Алуca AI - не справилась с заданием и не смогла ни определить тип шифра, ни расшифровать текст.

Claude - также не смог правильно выполнить задание. После одной подсказки нейросеть дала неверный ответ, а дальнейшее тестирование стало невозможным из-за ограничения бесплатного лимита запросов.

Perplexity AI - смогла определить тип шифра и расшифровать текст только после двух подсказок, когда вопрос был переформулирован и уточнён.

Microsoft Copilot - показал результат, сопоставимый с Perplexity AI: после двух подсказок нейросеть правильно определила шифр и величину сдвига.

Notion AI - справилась с заданием после одной подсказки, при условии уточнения используемого алфавита, и корректно расшифровала текст.

MiMo-V2 - не смогла правильно определить шифр с первой попытки. Нейросеть рассуждала в течение 3–4 минут, однако после ввода подсказки о величине сдвига перестала корректно реагировать на запросы, что сделало выполнение задания невозможным.

Сравнительная таблица результатов

<i>Нейросеть</i>	<i>Определение шифра</i>	<i>Подсказки</i>	<i>Время</i>
ChatGPT	Да	2	~46 сек
Gemini (быстрая)	Частично	5	—
Gemini (думающая)	Да	0–1	54 сек
DeepSeek	Нет	2	>2 мин
Алиса AI	Нет	—	—
Claude	Нет	1	—
Perplexity AI	Да	2	—
Copilot	Да	2	—
Notion AI	Да	1	—
MiMo-V2	Не	2	3–4 мин

Результаты выполнения задания «Разгадывание шифров»

<i>Нейросеть</i>	<i>Правильность</i>	<i>Логика объяснение</i>	<i>Полнота ответа</i>	<i>Структура</i>	<i>Время попытки</i>	<i>Итого /10</i>
ChatGPT	3	2	2	1	0	8

Нейросеть	Правильность	Правила объяснение	Логичность ответа	Поиск информации	Структура	Время попытки	Баллы /10
Gemini (думающая модель)	4	2	2	1	0	9	
Deep Seek	2	1	1	1	0	5	
Алиса AI	1	1	1	1	0	4	
Perplexity AI	3	2	1	1	0	7	
Microsoft Copilot	3	2	2	1	0	8	
Notion AI	4	2	2	1	0	9	
GigaChat	2	1	1	1	0	5	
MiM o-V2	2	1	1	1	0	5	

Выводы по разделу

По результатам данного этапа исследования можно сделать вывод, что задачи на разгадывание шифров оказались сложными для большинства нейросетей. Лишь часть моделей смогла корректно определить тип шифра и величину сдвига, причём в большинстве случаев это происходило **только после использования подсказок**.

Наиболее стабильные результаты показали ChatGPT, Gemini (думающая модель), Perplexity AI, Microsoft Copilot и Notion AI. При этом даже они не всегда справлялись с заданием с первой попытки.

Таким образом, можно заключить, что нейросети в целом **не являются надёжным инструментом для самостоятельного решения криптографических и логических задач**, особенно без дополнительного вмешательства пользователя. Это подтверждает необходимость критического подхода к использованию нейросетей в учебной деятельности.

3.2. Анализ текста и составление тезисного плана

Цель данного этапа исследования - оценить способность нейросетей структурировать текст, выделять ключевые идеи и формировать тезисный план с информативными заголовками.

Для эксперимента использовался фрагмент текста о необходимости охраны природы и экологическом развитии общества, включающий исторические примеры, современную экологическую проблематику и научную аргументацию.

Нейросетям было предложено:

1. Составить тезисный план текста.
2. Придумать 3 заголовка, отражающие смысл текста.

Результаты эксперимента

Нейросети показали различный уровень качества в выполнении задания.

Таблица результатов по 10-балльной шкале

Нейросеть	Правильность	Правильность	Логичность	Поиск	Структура	Время	Баллы /10
GPT Chat	3		2	2	1	1	9
Gemini (думающая)	4		2	2	1	1	10
Deep Seek	4		2	2	1	1	10
Alibaba AI	3		2	2	1	1	9
Perplexity AI	3		2	2	1	1	8
Copilot	4		2	2	1	1	10
Notion AI	4		2	2	1	1	9
GigaChat	3		1	1	1	1	7
MiM-V2	4		2	2	1	1	10

Примечания к оценке:

- **Правильность** - соответствие тезисов смыслу текста; выделение ключевых идей.

- **Логика** - последовательность, плавность переходов между тезисами.
- **Полнота** - охват всех важных аспектов текста.
- **Структура** - наличие заголовков и деление на смысловые блоки.
- **Время** - скорость выполнения задачи; нейросети, требовавшие подсказок или дополнительного времени, получили меньше баллов.

Качественный анализ

1. **ChatGPT**

- Плюсы: логично строит тезисы, выделяет основные идеи.
- Минусы: заголовки общие, без креативности.
- Баллы: 9/10

2. **Gemini (думающая модель)**

- Плюсы: хорошо структурирует текст, креативные заголовки, четкая логика.
- Минусы: мелкие детали упущены в некоторых тезисах.
- Баллы: 10/10

3. **DeepSeek**

- Плюсы: максимально полно отражает содержание, выразительные заголовки.
- Минусы: иногда избыточно длинные формулировки.
- Баллы: 10/10

4. **Алиса AI**

- Плюсы: выделяет основные тезисы, ясная структура.
- Минусы: заголовки менее информативные, пропущены некоторые нюансы.
- Баллы: 9/10

5. **Perplexity AI**

- Плюсы: логичная структура, кратко.
- Минусы: некоторые ключевые детали не выделены, заголовки простые.
- Баллы: 8/10

6. Microsoft Copilot

- Плюсы: полный и развернутый план, хорошие заголовки.
- Минусы: немного перегруженные тезисы.
- Баллы: 10/10

7. Notion AI

- Плюсы: четкая структура, выразительные заголовки.
- Минусы: не все детали текста включены.
- Баллы: 9/10

8. GigaChat

- Плюсы: план есть, базовые идеи выделены.
- Минусы: неполная логика, теряются детали, заголовки слишком простые.
- Баллы: 7/10

9. MiMo-V2

- Плюсы: полный охват, хорошие заголовки, логика соблюдена.
- Минусы: минимальные упущения, быстро формирует результат.
- Баллы: 10/10

Промежуточный вывод по главе

- **Лучшие результаты** показали Gemini (думающая модель), DeepSeek, Microsoft Copilot и MiMo-V2.
- **Хорошие, но менее креативные:** ChatGPT, Алиса AI, Notion AI.
- **Слабее остальных:** GigaChat, Perplexity AI (план неполный, заголовки упрощённые).

В целом нейросети демонстрируют высокую эффективность в структурировании текста и составлении тезисного плана, однако креативность заголовков и полнота охвата деталей зависят от модели. Балльная система позволяет наглядно сравнить эффективность и подготовить обоснованный рейтинг для итоговой главы.

3.3. Написание эссе на тему «Влияние музыки на успеваемость учащихся»

Цель данного этапа исследования - оценить способность нейросетей **создавать связные, аргументированные и структурированные тексты**, соответствующие заданной теме, с учётом научного подхода и доступного школьного стиля.

Нейросетям было предложено написать **эссе на заданную тему** и учесть:

- раскрытие положительных и отрицательных эффектов музыки;
- ссылки на исследования и достоверные факты;
- структуру (введение, основная часть, выводы);
- понятность изложения для школьников.

Таблица результатов по 10-балльной системе

	Нейр осеть	Правил ьность	Ло гика	По лнота	Стру ктура	В ремя	Ба ллы /10
GPT	Chat	3	2	2	3	1	9
ni	Gemini	4	2	2	4	1	10
Seek	Deep	4	2	2	4	1	10
a AI	Алис	3	2	2	3	1	9
exity AI	Perplexity	3	2	1	3	1	8
ot	Copilot	4	2	2	4	1	10
n AI	Notion AI	3	2	2	3	1	9

	Нейр	Правил	Ло	По	Стру	В	Ба
осеть	ьность	гика	лнота	ктура	ремя	ллы	/10
Chat	Giga	3	1	1	1	1	7
o-V2	MiM	4	2	2	1	1	10

Примечания к оценке:

- **Правильность** - соответствие теме, отсутствие фактических ошибок.
- **Логика** - плавность рассуждений, связность текста.
- **Полнота** - охват положительных и отрицательных эффектов, ссылки на исследования.
- **Структура** - наличие введения, основной части и заключения.
- **Время** - скорость создания эссе; балл снижается при необходимости повторных попыток или подсказок.

Качественный анализ

1. ChatGPT

- Плюсы: простая и понятная структура, корректные аргументы, раскрыты основные эффекты музыки.
- Минусы: нет конкретных цифр, мало ссылок на исследования.
- Баллы: 9/10

2. Gemini (думающая модель)

- Плюсы: полностью раскрывает тему, аргументы подкреплены исследованиями, структура чёткая.
- Минусы: иногда стиль немного академичен для школьника.
- Баллы: 10/10

3. DeepSeek

- Плюсы: подробное эссе, чётко выделены положительные и отрицательные эффекты, академичный стиль.
- Минусы: нет конкретных числовых данных, но по смыслу глубокое.

- Баллы: 10/10

4. **Алиса AI**

- Плюсы: развернутый текст, упомянуты дополнительные факты, указаны источники.
- Минусы: нет глубокого анализа, небольшие пропуски в структуре.
- Баллы: 9/10

5. **Perplexity AI**

- Плюсы: ссылки на исследования присутствуют, краткая и логичная подача.
- Минусы: неполная раскрытость темы, меньше практических примеров.
- Баллы: 8/10

6. **Microsoft Copilot**

- Плюсы: отлично структурированное и развернутое эссе, примеры и ссылки на исследования.
- Минусы: стиль местами слишком академичен для школьного эссе.
- Баллы: 10/10

7. **Notion AI**

- Плюсы: ясная структура, аргументы представлены логично, стиль доступен.
- Минусы: нет конкретных цифр, мало практических примеров.
- Баллы: 9/10

8. **GigaChat**

- Плюсы: базовая логика и структура эссе сохранены.
- Минусы: неполнота аргументов, простые формулировки, мало ссылок.
- Баллы: 7/10

9. **MiMo-V2**

- Плюсы: полное эссе, хорошие аргументы, структура сохранена.
- Минусы: минимальные упущения, стиль академичный.
- Баллы: 10/10

Промежуточный вывод по главе

- **Лучшие результаты:** Gemini (думающая), DeepSeek, Microsoft Copilot и MiMo-V2 - 10/10.
- **Хорошие результаты:** ChatGPT, Алиса AI, Notion AI - 9/10.
- **Слабее остальных:** Perplexity AI и GigaChat - 8 и 7 баллов соответственно.

Нейросети, которые получили максимальные баллы, продемонстрировали **наибольшую полноту раскрытия темы, логическую структуру и аргументированность**, а также справились с заданием быстро, без значительного времени на исправления и подсказки.

3.4. Решение математических задач

Цель данного этапа исследования - оценить способность нейросетей **решать математические задачи с точным учётом области допустимых значений (ОДЗ), логикой и корректностью вычислений.**

Для эксперимента было выбрано задание:

$$(x+3) \cdot (x^2-3x) > x^2-x-2$$

Нейросетям предлагалось:

- определить ОДЗ;
- решить неравенство;
- дать обоснование решения;
- при необходимости уточнить промежутки и ограничения.

Таблица результатов по 10-балльной системе

Нейросеть	Правильность	Правила	Логика	Полнота	Структура	Время	Баллы /10
ChatGPT		3	2	2	1	0	8

	Нейр	Правил	Ло	По	Стру	В	Ба
осеть	ьность	гика	лнота	ктура	ремя	ллы	/10
ni	Gemini	3	2	2	1	1	9
Seek	Deep	3	2	2	1	1	9
a AI	Алис	2	1	2	1	1	7
exity AI	Perplexity	2	1	1	1	1	6
rosoft	Microsoft	3	2	2	1	0	8
Copilot							
n AI	Notion	3	2	2	1	0	8
Chat	GigaChat	1	1	1	1	1	5
o-V2	MiM-V2	3	2	2	1	0	8

Примечания к оценке:

- **Правильность** - корректное решение, точное определение промежутков.
- **Логика** - последовательность рассуждений и объяснений.
- **Полнота** - учёт всех деталей, включая ОДЗ.
- **Структура** - разбиение решения на шаги, ясность изложения.
- **Время** - скорость получения корректного решения без необходимости подсказок.

Качественный анализ

1. **ChatGPT**

- Плюсы: правильно решил неравенство с учётом большинства условий, объяснение последовательное.
- Минусы: сложное решение, ответ не полностью совпадает с точной областью ОДЗ.
- Баллы: 8/10

2. **Gemini (думающая модель)**

- Плюсы: корректное решение, учитывает ОДЗ, логика чёткая.
- Минусы: немного громоздко изложено, но верно.
- Баллы: 9/10

3. **DeepSeek**

- Плюсы: правильный ответ, логика решения последовательная.
- Минусы: небольшие лишние рассуждения, но без ошибок.
- Баллы: 9/10

4. **Алиса AI**

- Плюсы: частично правильное решение, попытка учёта ОДЗ.
- Минусы: ошибки в интервалах, неполное объяснение.
- Баллы: 7/10

5. **Perplexity AI**

- Плюсы: решение присутствует.
- Минусы: неверное определение промежутков, ОДЗ не полностью учтена.
- Баллы: 6/10

6. **Microsoft Copilot**

- Плюсы: верное решение для одного интервала, структура ясная.
- Минусы: не рассматривает полностью все промежутки.
- Баллы: 8/10

7. **Notion AI**

- Плюсы: правильно решает задачу, логично объясняет.

- Минусы: промежутки указаны частично.
- Баллы: 8/10

8. GigaChat

- Плюсы: попытка решения есть.
- Минусы: некорректный ответ, ОДЗ не учтена.
- Баллы: 5/10

9. MiMo-V2

- Плюсы: правильно решает, логика последовательная, структура хорошая.
- Минусы: минимальные недочёты в оформлении.
- Баллы: 8/10

Промежуточный вывод по главе

- **Наиболее точные результаты:** Gemini (думающая модель) и DeepSeek - 9/10.
- **Хорошие результаты:** ChatGPT, Microsoft Copilot, Notion AI, MiMo-V2 - 8/10.
- **Средний уровень:** Алиса AI - 7/10.
- **Слабый результат:** Perplexity AI и GigaChat - 6 и 5 баллов соответственно.

В целом нейросети демонстрируют ограниченную способность решать математические задачи с полным учётом ОДЗ, даже если они успешно справляются с текстовыми и эссе заданиями. Балльная система позволяет объективно ранжировать эффективность каждой модели.

3.5. Исторические факты

Цель этапа - оценить способность нейросетей **правильно отвечать на вопросы по истории, давать точные даты, имена и контекст событий**, а также указывать источники при необходимости.

Задания для нейросетей:

1. Когда и где состоялся Второй континентальный конгресс?
2. Когда была принята Декларация независимости?
3. Кто был автором Декларации независимости?
4. Когда Франция объявила войну Англии?
5. Почему указ о запрете заселять земли за Аллеганским хребтом вызвал массовое недовольство колонистов?

Таблица результатов по 10-балльной системе

	Нейр	Правил	Ло	По	Стру	В	Ба
осеть	ьность	гика	лнота	ктура	ремя	ллы	/10
GPT	Chat	3	2	2	1	0	8
ni	Gemini	4	2	2	1	0	9
Seek	Deep	3	2	2	1	1	8
a AI	Алис	4	2	2	1	0	9
exity AI	Perplexity	3	1	1	1	1	7
rosoft Copilot	Microsoft	3	2	2	1	0	8
n AI	Notion	4	2	2	1	0	9
Chat	GigaChat	3	1	2	1	0	7

	Нейр	Правил	Ло	По	Стру	В	Ба
осеть	ьность	гика	лнота	ктура	ремя	ллы	/10
MiM							
o-V2	3		2	2	1	0	8

Примечания:

- Основная сложность была в вопросе 4: «Когда Франция объявила войну Англии». Этот вопрос спорный, потому что точная дата требует уточнения о союзе с США.

Качественный анализ

1. **ChatGPT:** правильные ответы на все вопросы, кроме спорного 4-го, подробные объяснения - 8/10
2. **Gemini:** точные даты и авторы, полный контекст, логика ясна - 9/10
3. **DeepSeek:** верно отвечает, но упрощает объяснение спорного вопроса - 8/10
4. **Алиса AI:** правильные даты и авторы, краткие пояснения, ссылки на факты - 9/10
5. **Perplexity AI:** ответы частично правильные, 4-й вопрос выполнен некорректно - 7/10
6. **Microsoft Copilot:** правильные даты, но 4-й вопрос не до конца раскрыт - 8/10
7. **Notion AI:** точные ответы, ясная логика, кратко - 9/10
8. **GigaChat:** простые ответы, логика слабее - 7/10
9. **MiMo-V2:** правильные ответы, ясное объяснение, 4-й вопрос частично - 8/10

Вывод по истории:

- Лучшие результаты: Gemini, Алиса AI, Notion AI - 9/10
- Средний уровень: ChatGPT, DeepSeek, Microsoft Copilot, MiMo-V2 - 8/10
- Слабее остальных: Perplexity AI, GigaChat - 7/10

Вопросы с неопределённым контекстом выявляют ограничения нейросетей в исторической точности и необходимости уточнений.

3.6. Генерация кода (игра «Змейка»)

Цель этапа - оценить способность нейросетей **создавать рабочий код** для вебигры на HTML+CSS+JavaScript, учитывая:

- работоспособность игры;
- интерфейс и удобство управления;
- корректность и читаемость кода;
- скорость написания.

Таблица результатов по 10-балльной системе

	Нейросеть	Работоспособность	Логичность кода	Поинтерфейс	Интерфейс	Время	Баллы /10
	Chat GPT	4	2	2	2	0	10
	Gemini	4	2	2	2	0	10
	Deep Seek	4	2	2	2	0	10
	Алиса AI	0	0	0	0	0	0
	Perplexity AI	2	1	1	1	0	5
	Microsoft Copilot	4	2	2	2	0	10

	Ней росеть	Работоспо собность	Ло гика кода	По лнота	Инте рфейс	В ремя	Б аллы /10
Noti on AI		3	2	2	1	0	8
Giga Chat		1	1	1	0	0	3
MiM o-V2		4	2	2	1	0	9

Качественный анализ

1. **ChatGPT, Gemini, DeepSeek, Copilot:**

- Плюсы: игра запускается, багов нет, управление удобно, код читаемый - 10/10

2. **Notion AI:**

- Плюсы: код рабочий, игра запускается, управление понятно
- Минусы: интерфейс простой, немного быстрая змейка - 8/10

3. **MiMo-V2:**

- Плюсы: рабочий код, быстрый результат
- Минусы: интерфейс простой, управление чуть менее удобное - 9/10

4. **Perplexity AI:**

- Плюсы: код генерируется, игра запускается
- Минусы: ошибки в логике игры, сразу поражение - 5/10

5. **GigaChat:**

- Плюсы: код есть
- Минусы: игра сразу ломается, интерфейс неудобный - 3/10

6. **Алиса AI:**

- Не смогла сгенерировать рабочий код - 0/10

Вывод по коду:

- Лидеры: ChatGPT, Gemini, DeepSeek, Copilot - идеально для школьного проекта.
- Средний уровень: MiMo-V2, Notion AI - небольшие недочёты в интерфейсе.
- Слабые: Perplexity AI, GigaChat, Алиса AI - баги, нерабочий код.

Теперь у нас есть **полные данные по всем практическим заданиям (3.1–3.6) с балльной системой**, и мы можем **составить итоговую сводную таблицу и сделать финальный рейтинг нейросетей с сильными сторонами каждой**.

4. Итоговая оценка и рейтинг нейросетей

Цель данного раздела - подвести итоги исследования, определить, **какая нейросеть показала наибольшую эффективность** во всех типах заданий, и **выявить сильные и слабые стороны каждой модели**.

4.1. Сводная таблица баллов

Нейросеть	Шифры	Шекст	Тссе	Эматика	Математика	История	История	Кто	Итого
ChatGPT	8		9	9	8	8	0	12	5
Gemini (думающая)	9	0	10	11	9	9	0	7	5
DeepSeek	9	0	10	11	9	8	0	6	5
Алиса AI	0		9	9	7	9		04	3

	Ней	Ш	Т	Э	Мате	Ист	К	И
росеть	ифры	екст	ссе	матика	ория	од	того	
Perpl								4
exity AI	7		8	8	6	7	5	1
Micr								
rosoft	8		1	1	8	8	1	5
Copilot		0	0			0	4	
Notio								
n AI	9		9	9	8	9	8	2
Giga								
Chat	7		7	7	5	7	3	6
MiM								
o-V2	8		1	1	8	8	9	3

Claude не был включён в итоговую оценку из-за ограничений бесплатной версии.

4.2. Итоговый рейтинг

1. Gemini (думающая модель) - 57 баллов

- **Сильные стороны:** универсальность, точность в тексте и эссе, корректное решение задач, быстрая и качественная генерация кода.
- **Слабые стороны:** небольшая громоздкость в изложении математических решений.

2. DeepSeek - 56 баллов

- **Сильные стороны:** полный охват текста, эссе, точная генерация кода, логика и последовательность рассуждений.
- **Слабые стороны:** небольшие упрощения в исторических данных.

3. Microsoft Copilot - 54 балла

- **Сильные стороны:** идеально подходит для кода и эссе, стабильный результат во всех заданиях.

- **Слабые стороны:** математические решения иногда требуют уточнения.
4. **MiMo-V2** - 53 балла
- **Сильные стороны:** точность текста и эссе, рабочий код, логика решения задач.
 - **Слабые стороны:** интерфейс кода и скорость решения шифров оставляют желать лучшего.
5. **ChatGPT** - 52 балла
- **Сильные стороны:** стабильность в тексте и коде, логика, структура, доступность языка.
 - **Слабые стороны:** иногда ошибки в математике и шифрах, требуется подсказка.
6. **Notion AI** - 52 балла
- **Сильные стороны:** ясная структура текстов, хорошие заголовки, сильна в истории и эссе.
 - **Слабые стороны:** код и шифры чуть уступают лидерам.
7. **Perplexity AI** - 41 балл
- **Сильные стороны:** ссылки на источники, краткость.
 - **Слабые стороны:** ошибки в математике, коде и некоторых заданиях, требует подсказок.
8. **GigaChat** - 36 баллов
- **Сильные стороны:** минимально выполняет все задания.
 - **Слабые стороны:** слабая логика, ошибки в коде, неполные ответы.
9. **Алиса AI** - 34 балла
- **Сильные стороны:** хорошо справляется с историей и эссе.
 - **Слабые стороны:** не справляется с кодом и шифрами, ограниченный функционал.

4.3. Общие выводы

1. **Лучшие нейросети для комплексного использования:** Gemini (думающая модель), DeepSeek, Microsoft Copilot - демонстрируют **высокую универсальность** и качество во всех типах заданий.
2. **Лучшие для анализа текста и написании эссе:** MiMo-V2, ChatGPT, Notion AI - сильны в **структурировании текста, эссе и тезисных планах**, менее эффективны в коде и математике.
3. **Ограниченные нейросети:** Perplexity AI, GigaChat и Алиса AI - показывают неплохие результаты в отдельных заданиях, но **не универсальны**.
4. **Балльная система** показала свою эффективность, так как позволила **объективно оценивать нейросети** и выявить сильные стороны каждой модели.

Итог: **Gemini (думающая модель)** является лучшей универсальной нейросетью по совокупности всех критериев, в то время как другие модели имеют узкую специализацию и слабые стороны, выявленные в эксперименте. Результаты могут быть использованы учащимися при выборе нейросети для конкретных учебных задач.

Заключение

В ходе данной научно-исследовательской работы было проведено сравнительное исследование современных нейросетей с целью оценки их эффективности при выполнении различных типов заданий. В качестве объектов исследования были выбраны наиболее популярные нейросети, широко используемые в образовательной и практической деятельности.

Для достижения поставленной цели была разработана и применена **единая 10-балльная система оценки**, учитывающая правильность ответа, логичность рассуждений, полноту, структурированность и время выполнения заданий.

Такой подход позволил обеспечить объективность сравнения и свести субъективность оценки к минимуму.

В рамках практической части нейросети выполняли задания по разгадыванию шифров, анализу текста, написанию эссе, решению математических задач, работе с историческими фактами и генерации программного кода. Результаты показали, что **ни одна нейросеть не является абсолютно универсальной**, однако некоторые модели демонстрируют стабильно высокие показатели в большинстве заданий.

По итогам суммарного подсчёта баллов было установлено, что наибольшую эффективность показала **Gemini (думающая модель)**, продемонстрировав высокий уровень универсальности и качества ответов. **DeepSeek** и **Microsoft Copilot** также заняли лидирующие позиции, особенно в задачах, связанных с программированием и работой с текстом. **ChatGPT**, **Notion AI** и **MiMo-V2** показали хорошие результаты в гуманитарных и текстовых заданиях, однако уступили лидерам в математике и шифрах. **Perplexity AI**, **GigaChat** и **Алиса AI** оказались менее универсальными и показали нестабильные результаты в ряде заданий.

Проведённое исследование подтвердило, что использование балльной системы является эффективным инструментом для анализа и сравнения нейросетей. Полученные результаты могут быть использованы учащимися и педагогами для выбора нейросети в зависимости от поставленных задач - учебных, творческих или технических.

Таким образом, поставленные в работе цель и задачи были полностью достигнуты, а гипотеза о различной специализации и эффективности нейросетей нашла своё подтверждение.

Список литературы

- 1.Алдабек Н. Всемирная история. Учебник для 8 класса. - 2018.

2.Шабунин М.И., Прокофьев А.А., Олейник Т.А., Соколова Т.В.
Алгебра и начала математического анализа. Профильный уровень.
- М.: БИНОМ, 2009.

3.Статья: «Сравнение нейросетей» // Habr.

URL: <https://habr.com/ru/companies/bothub/articles/893128/>

(дата обращения: 2025).

4. Брюс Шнайер Прикладная криптография 1994.

5. Русский язык Сабитова З.К. ОГН общественно-гуманитарное
направление учебник для 10 класса-2009.

Literature list

1.Aldabek H. World history. Textbook for the 8th grade. - 2018.

2.Shabonin M.I., Prokofiev A.A., Oleinik T.A., Sokolova T.V.
Algebra and start of mathematical analysis. Profile level.
-М.: BINON, 2009.

3.Article: «Comparison of neural networks» // Habr.

URL: <https://habr.com/ru/companies/bothub/articles/893128/>

(date of circulation: 2025).

4. Bruce Schneider Applied Cryptography 1994.

5. Russian language Sabitova Z.K. OGN public-humanitarian direction
textbook for class 10-2009.

Приложения:

Приложение А

Примеры заданий и ответов нейросетей по разгадыванию шифров

В данном приложении приведены примеры заданий по разгадыванию шифров и фрагменты ответов нейросетей, использованные в ходе исследования.

Задание:

Сдвиг вправо на 5 букв. Расшифруй предложение:
«Те цйзуитдэтно ийтб пурныйцчжу лнчйрйо ж цхетй фхйжернру ме ижеиыечб снррнутуж».

Пример ответа ChatGPT:

Нейросеть предположила использование шифра Цезаря, однако на первой попытке дала неточный результат. После уточнения типа шифра и направления сдвига был получен корректный вариант расшифровки.

Пример ответа Gemini (думающая модель):

Нейросеть самостоятельно определила тип шифра как шифр Цезаря и корректно установила величину сдвига, после чего дала правильный расшифрованный текст.

Приведённые примеры иллюстрируют различия в скорости анализа, количестве попыток и способности нейросетей самостоятельно определять тип шифра.

Приложение Б

Таблицы балльной оценки нейросетей

В данном приложении представлены сводные таблицы балльной оценки нейросетей по всем видам заданий, выполненных в рамках исследования.

Оценивание проводилось по 10-балльной системе с учётом следующих критериев:

- правильность ответа;
- логика и обоснование решения;
- полнота;
- структурированность;

- время получения результата.

В таблицах отражены результаты по следующим заданиям:

- разгадывание шифров;
- анализ текста и составление тезисного плана;
- написание эссе;
- решение математических задач;
- исторические факты;
- генерация программного кода.

Данные таблицы использовались для составления итогового рейтинга нейросетей и анализа их сильных и слабых сторон.

Приложение В

Примеры эссе, сгенерированных нейросетями

В данном приложении представлен пример эссе, созданного одной из нейросетей в ходе выполнения задания.

Тема

эссе:

Влияние музыки на успеваемость учащихся.

Эссе содержит рассуждения о положительном и отрицательном влиянии музыки на концентрацию внимания, эмоциональное состояние и когнитивные способности учащихся, а также вывод о необходимости индивидуального подхода при использовании музыки во время обучения.

Пример эссе был использован для оценки таких параметров, как логичность изложения, глубина аргументации, структура текста и научная обоснованность.

Приложение Г

Пример программного кода веб-игры «Змейка»

В данном приложении представлен пример программного кода веб-игры «Змейка», сгенерированного нейросетью в рамках практической части исследования.

Код реализован в виде одного HTML-файла и включает:

- HTML-разметку;
- стили CSS;
- сценарий на JavaScript.

Игра поддерживает управление с помощью клавиш-стрелок и клавиш WASD, корректно запускается в браузере и позволяет управлять игровым процессом.

Данный пример использовался для оценки способности нейросетей генерировать рабочий и логически выстроенный программный код.