



Как запустить эффективный ИИ и не потратить лишнего?



ВАЛЕНТИН КАСЬКОВ

СІО ГК ССТ

Эксперт в области

- ✓ Искусственного интеллекта
- ✓ Цифровых трендов
- ✓ Автоматизации бизнес-процессов

Специализация

- ✓ Разработка ПО для инжиниринговых компаний
- ✓ Создание «цифровых инженеров» — систем расчёта и принятия решений в инженерных проектах

Преимущества внедрения ИИ в цифровые продукты



Предиктивная аналитика



Автоматизация отчетности
и аналитики



Сокращение трудозатрат

+5%

и более повышение
рентабельности
в краткосрочном периоде

Бизнес-заказчик не является экспертом в сфере ИИ

Заказчик не обладает необходимыми компетенциями, чтобы оценить качество внедренной системы

- 1 Внедрен ли ИИ?
- 2 Как оценить работоспособность и качество внедренной LLM?

1 **Как заказчику
определить
внедрен ли ИИ?**

Критерий оценки

Система решает поставленные задачи без участия людей

2

Как оценить работоспособность и качество внедренной LLM?

Критерий оценки

- ГОСТ Р59898-2021
- Методика «Сбера»
- Оценка качества ответов с помощью каскадного тестирования

Аудит внедрения ИИ ГОСТ Р 59898-2021

Стандарт распространяется на все системы, использующие различные методы ИИ, для решения конкретных практически значимых задач



Функциональность

- функциональные возможности
- степень автоматизации
- способность к взаимодействию
уровень производительности
- простота внедрения,
использования
- практичность



Надежность

- стабильность,
- устойчивость
к ошибке



Безопасность

- конфиденциальность
- подлинность
- приватность

Аудит внедрения ИИ

Методика «Сбера»



Скорость

среднее время выполнения клиентского пути



Качество

удовлетворенность клиентов цифровым продуктом



Объективность

доля решений, принятых без участия человека



Персонализация

доля продуктов/услуг, параметры которых подобраны под пользователя



Экономическая эффективность

затраты на единицу продукции

Каскадное тестирование LLM

Верификация

Производится верификация и проверка результатов, полученных предыдущими нейросетями

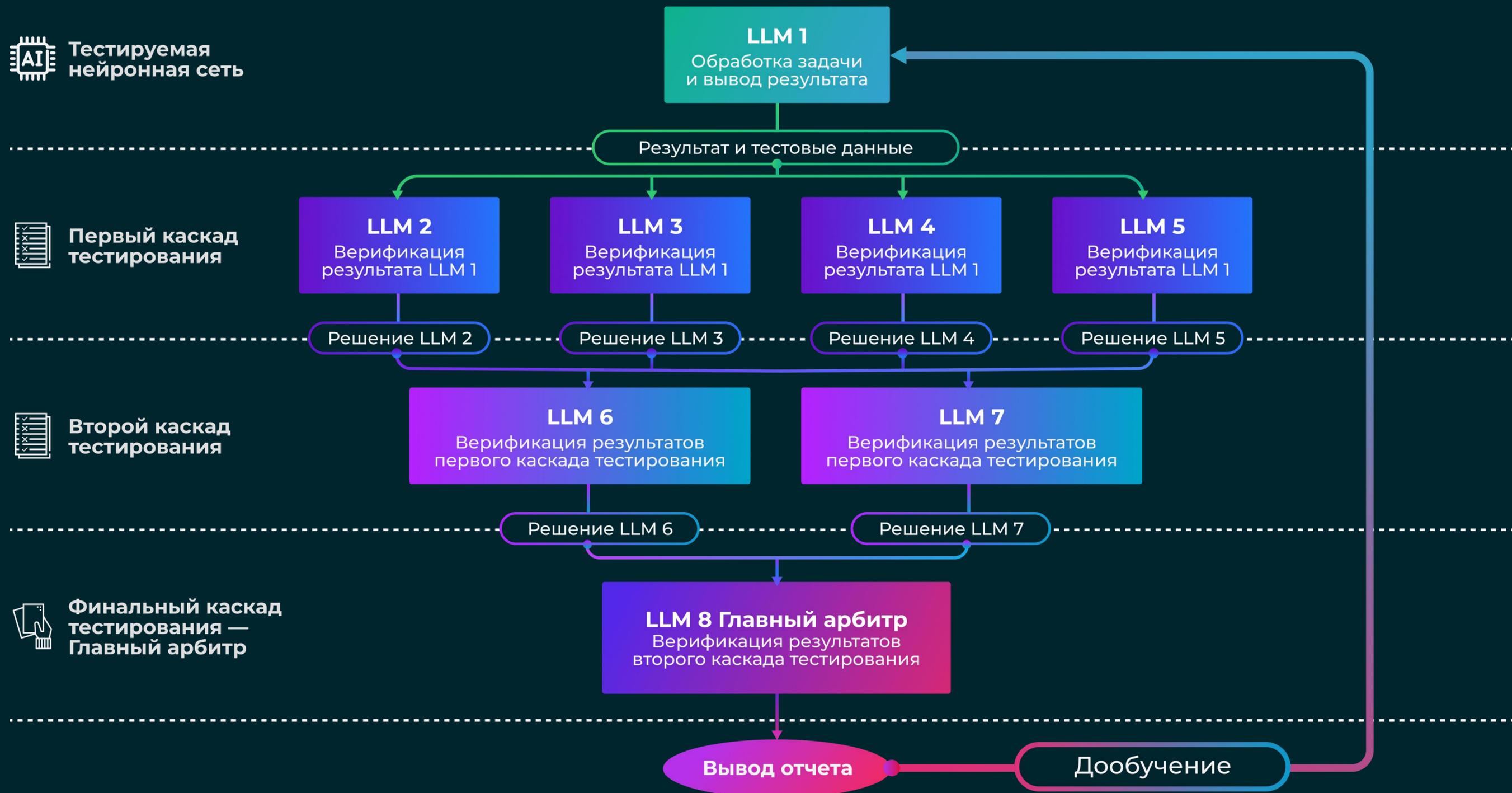
Что хорошо?

Что нужно улучшить?

Рекомендации

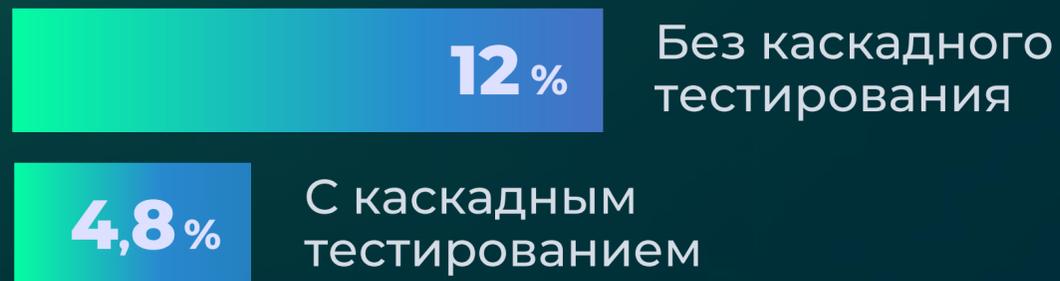
В результате верификации производится генерация обучающего контекста и дообучение проверяемых нейросетей

Процесс каскадного тестирования



Преимущества каскадного тестирования

Процент ошибочных ответов



60%

общий процент снижения аномалий в ответах

Подготовка к тестированию

Задача

Протестировать точность ответов LLM с помощью каскадного построения LLM

Цель

- ✓ Снизить количество аномалий/галлюцинаций в ответах LLM
- ✓ Автоматизировать процесс дообучения LLM

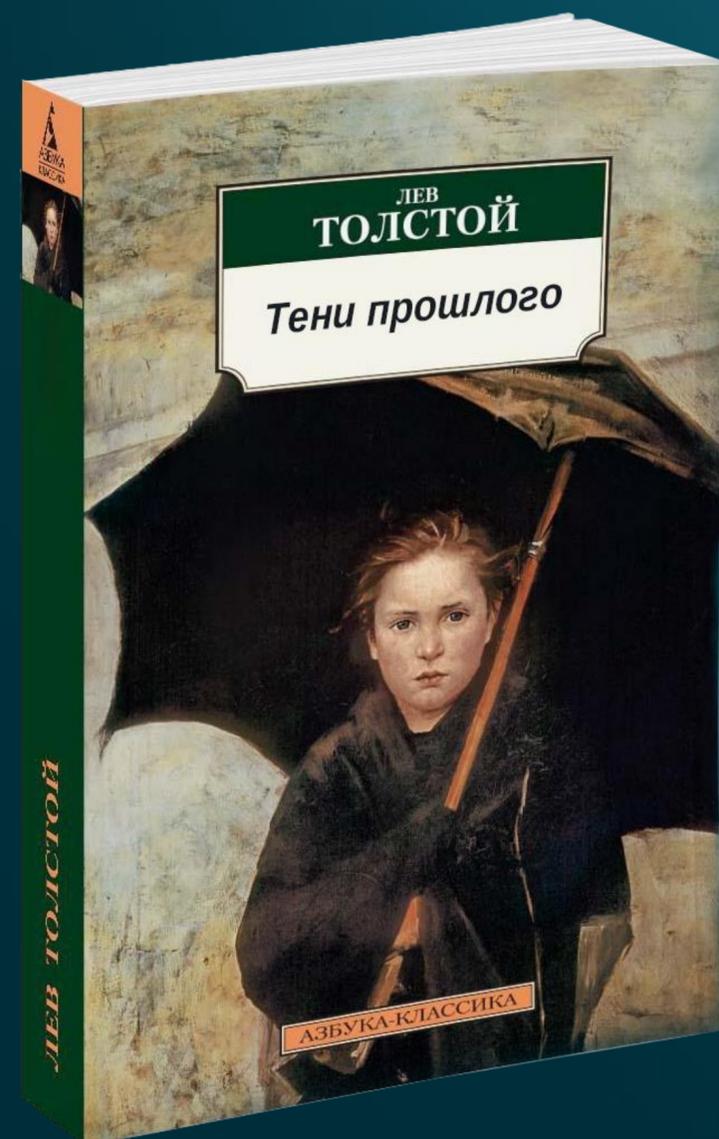
Case 1

Каковы основные темы и сюжет романа «Тени прошлого» Льва Толстого?



Ответ тестируемой LLM:

Роман «Тени прошлого» Льва Толстого основывается на истории любви между главными героями – молодым аристократом Николаем Ивановичем и простой деревенской девушкой Марьей Васильевной. Они влюбляются друг в друга, несмотря на разницу в социальном статусе и противодействие со стороны окружающих...



Case 1

Каковы основные темы и сюжет романа «Тени прошлого» Льва Толстого?



Ответ тестируемой LLM:

Роман «Тени прошлого» Льва Толстого основывается на истории любви между главными героями...

1

Решение LLM первого каскада:

Результаты не точны, требуется исправление и доработка ответа

2

Решение LLM второго каскада:

Роман «Тени прошлого» Льва Толстого не существует. Требуется доработка



Case 1

Каковы основные темы и сюжет романа «Тени прошлого» Льва Толстого?



Ответ тестируемой LLM:

Роман «Тени прошлого» Льва Толстого основывается на истории любви между главными героями...



Решение LLM первого каскада:

Результаты не точны, требуется исправление и доработка ответа



Решение LLM второго каскада:

Роман «Тени прошлого» Льва Толстого не существует. Требуется доработка



Финальный каскад

Верификация результатов

Тестируемая LLM: Ответ содержит описание романа «Тени прошлого» Льва Толстого, которого на самом деле не существует



LLM первого каскада: Ответ корректно указывает на ошибку

LLM второго каскада: Ответ корректно указывает на ошибку

Рекомендации

Ответ тестируемой LLM требует доработки, чтобы избежать предоставления неверной информации. Важно проверять наличие произведения в библиографии автора



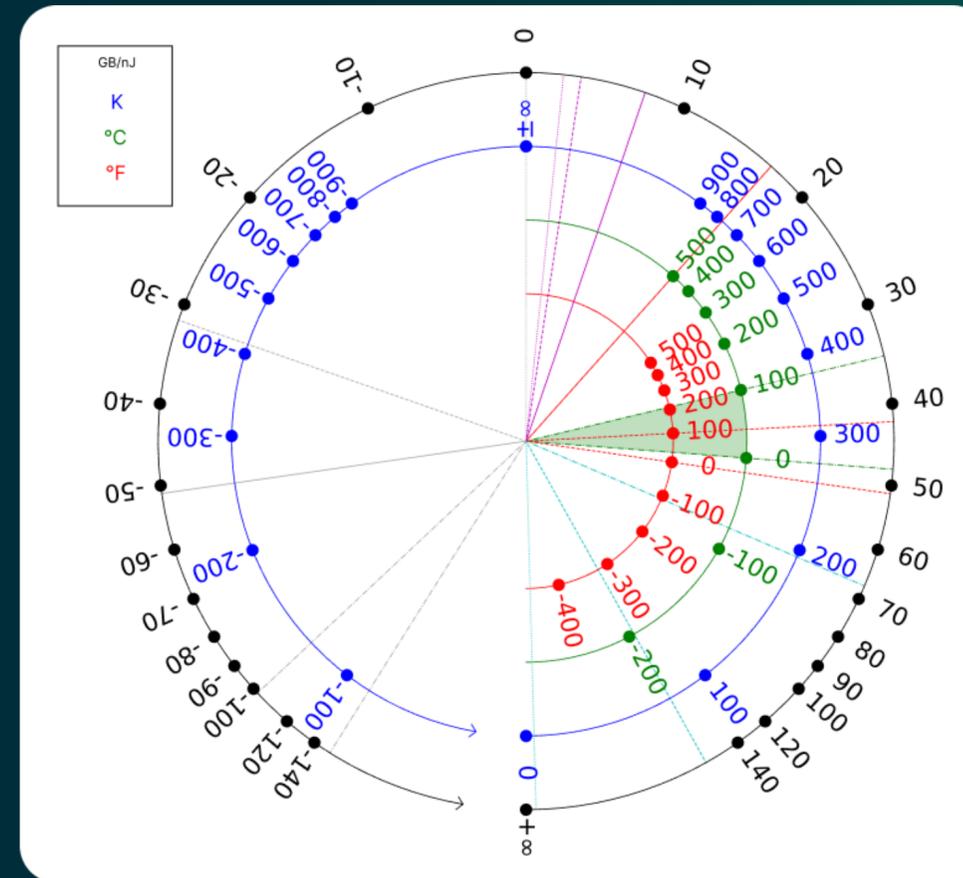
Case 2

Объясните теорию «инверсной термодинамики» и её применение в современной физике



Ответ тестируемой LLM:

Инверсная термодинамика - это область физики, которая изучает процессы, происходящие в системах, находящихся в нестандартных термодинамических условиях, таких как отрицательная температура или давление ...



Источник: <https://en.wikipedia.org/wiki/>

Case 2

Объясните теорию «инверсной термодинамики» и её применение в современной физике



Ответ тестируемой LLM:

Инверсная термодинамика - это область физики, которая изучает процессы, происходящие в системах...



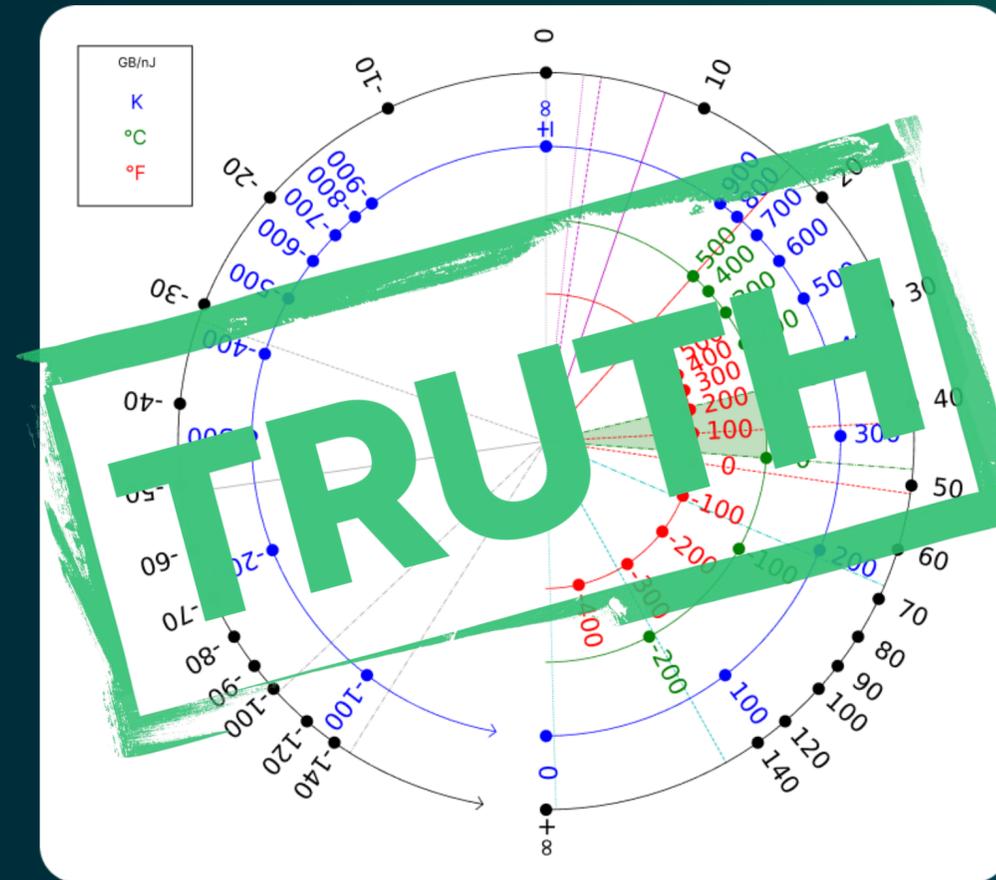
Решение LLM первого каскада:

Ответ LLM 1 точен и подробен



Решение LLM второго каскада:

Ответ LLM 1 точен, однако термин «инверсная термодинамика» не является общепринятым



Источник: <https://en.wikipedia.org/wiki/>

Case 2

Объясните теорию «инверсной термодинамики» и её применение в современной физике



Ответ тестируемой LLM:

Инверсная термодинамика - это область физики, которая изучает процессы, происходящие в системах...



Решение LLM первого каскада:

Ответ LLM 1 точен и подробен



Решение LLM второго каскада:

Ответ LLM 1 точен, однако термин «инверсная термодинамика» не является общепринятым



Финальный каскад

Верификация результатов

Тестируемая LLM: Термин «инверсная термодинамика» не является общепринятым в научной литературе

LLM первого каскада: Ответ LLM не точен

LLM второго каскада: Ответ LLM предлагает более точный термин «отрицательная температура»

Рекомендации

Ответ тестируемой LLM соответствует базовому уровню понимания, но требует уточнений



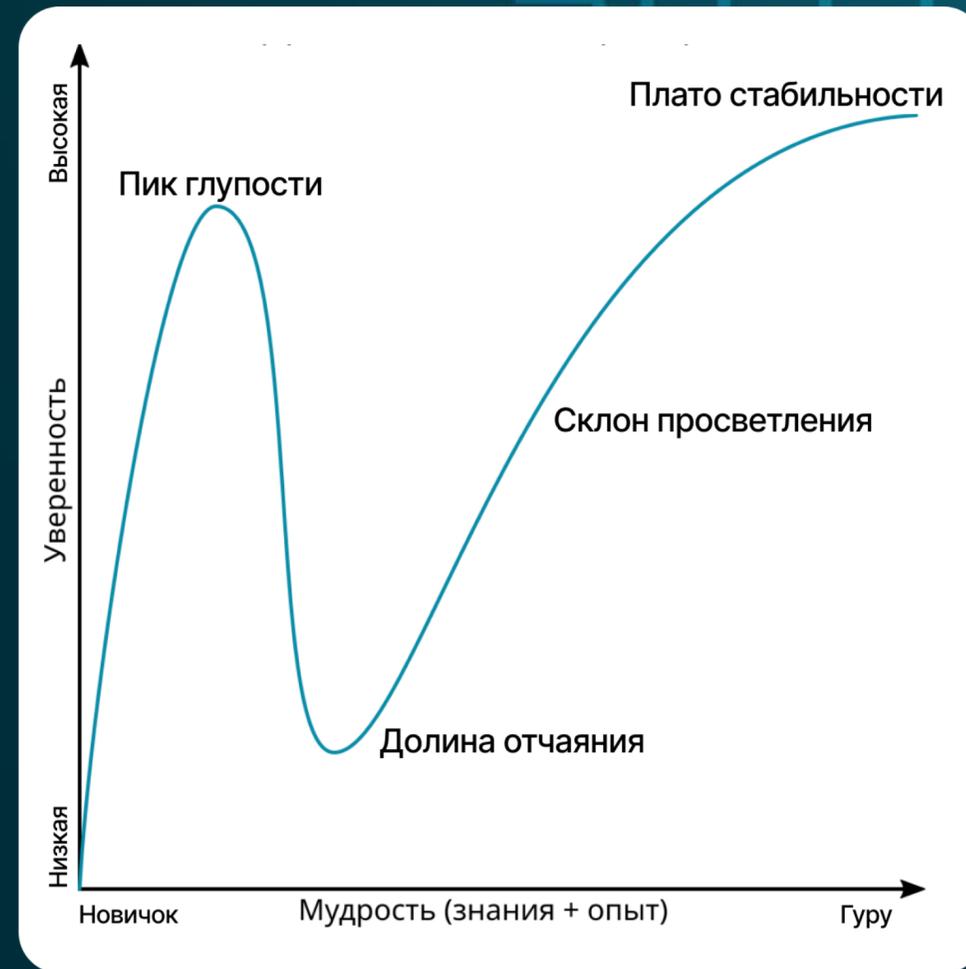
Case 3

Что такое эффект Даннинга-Крюгера?



Ответ LLM 1:

Эффект Даннинга-Крюгера - это когнитивное искажение, при котором люди с низким уровнем компетенции в определенной области часто недооценивают свои способности и знания, в то время как люди с высоким уровнем компетенции могут недооценивать свои способности из-за предположения, что другие люди также обладают такими же знаниями и умениями ...



Источник: <https://en.wikipedia.org/wiki/>

Case 3

Что такое эффект Даннинга-Крюгера?



Ответ тестируемой LLM:

Эффект Даннинга-Крюгера - это когнитивное искажение, при котором люди с низким уровнем компетенции в определенной области часто ...



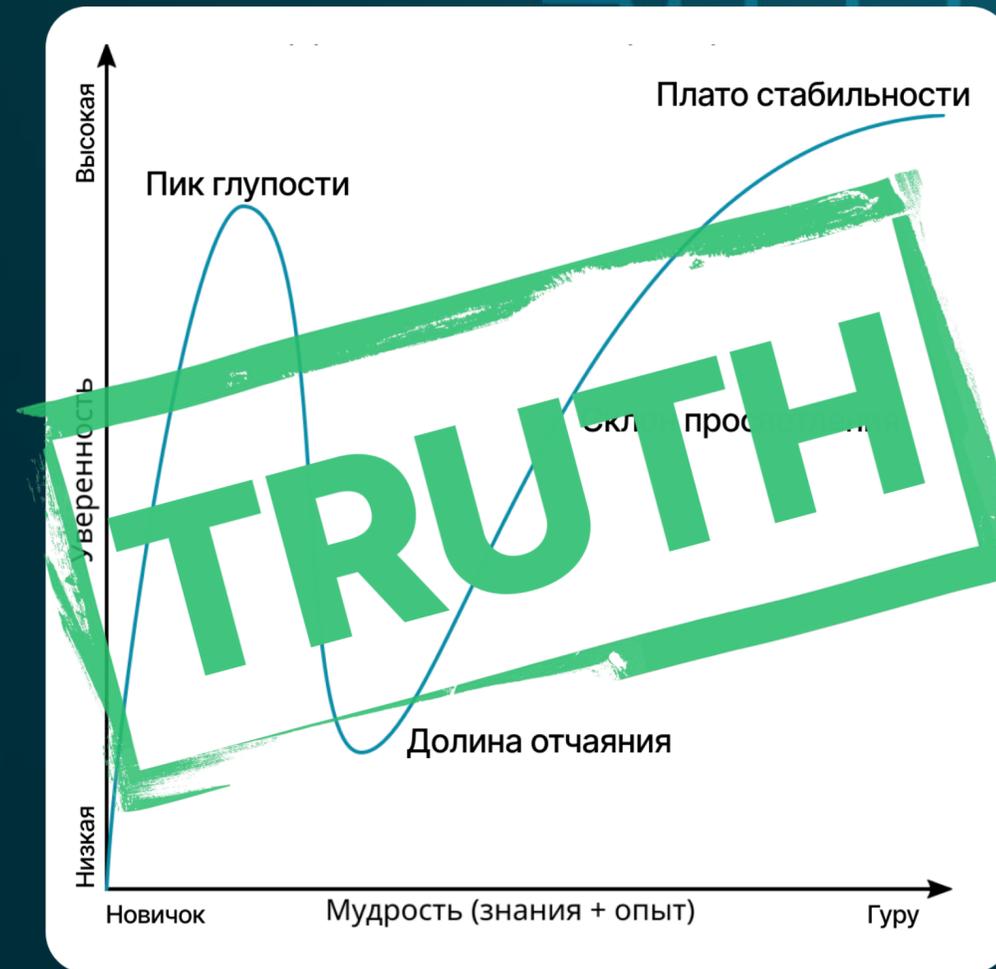
Решение LLM первого каскада:

Ответ точен и подробен



Решение LLM второго каскада:

Ответ точен и подробен



Источник: <https://en.wikipedia.org/wiki/>

Case 3

Что такое эффект Даннинга-Крюгера?



Ответ тестируемой LLM:

Эффект Даннинга-Крюгера - это когнитивное искажение, при котором люди с низким уровнем компетенции в определенной области часто ...



Решение LLM первого каскада:

Ответ точен и подробен



Решение LLM второго каскада:

Ответ точен и подробен



Финальный каскад

Верификация результатов

Тестируемая LLM: Ответ точно описывает эффект Даннинга-Крюгера как когнитивное искажение ...



LLM первого каскада:
Ответ точен и подробен

LLM второго каскада:
Ответ точен и подробен



Рекомендации

Ответ тестируемой LLM точен и подробен и не требует доработки или исправления

Команда **НИОКР**



Валентин Каськов

Эксперт по ИИ, цифровым трендам и оптимизации бизнес-процессов



Сергей Червяков

Эксперт по ИИ и цифровым продуктам

