



Мати

На пути к современному
образовательному процессу

Мати



Сидим
не рыпаемся



AI Challenge



ЦЕНТРАЛЬНЫЙ
УНИВЕРСИТЕТ

Если вы студент и выбираете, на чем сосредоточиться, выберите **МАТЕМАТИКУ**. Она научит вас неуклонно полагаться на собственный мозг, мыслить логически, разбирать проблемы и решать их шаг за шагом в правильном порядке. Это ключевой навык, который вам понадобится для создания компаний и управления проектами

— Павел Дуров



БОЛЬ И МОТИВАЦИЯ

Проверка математических задач требует большого количества времени и экспертизы, что затрудняет массовое обучение и снижает качество обратной связи для студентов.

5-12 минут

Занимает проверка
одной задачи вручную

30+ часов

На одно домашнее
задание в группе из 200
студентов



Студенты

Получают позднюю
обратную связь



Преподаватели и ассистенты

Перегружены рутиной

Почему это важно?

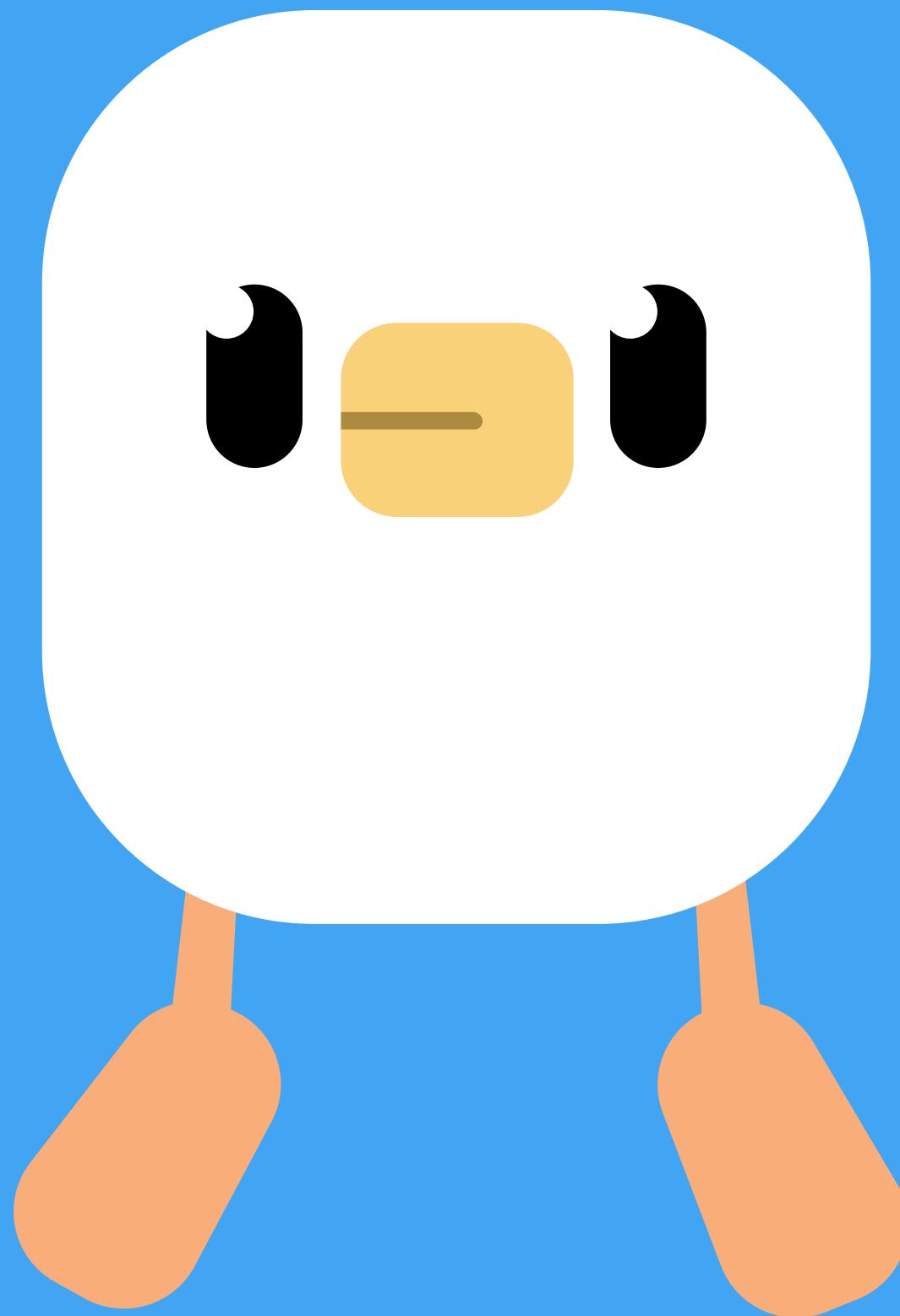
В IT- и инженерных направлениях наблюдается ежегодный рост числа студентов на **6–8%**.

Один преподаватель в техническом университете в среднем проверяет **150–250 работ** за неделю в период сессии.

75% преподавателей в СНГ жалуются на нехватку времени для проверки заданий (НИУ ВШЭ, 2023).

Наше решение

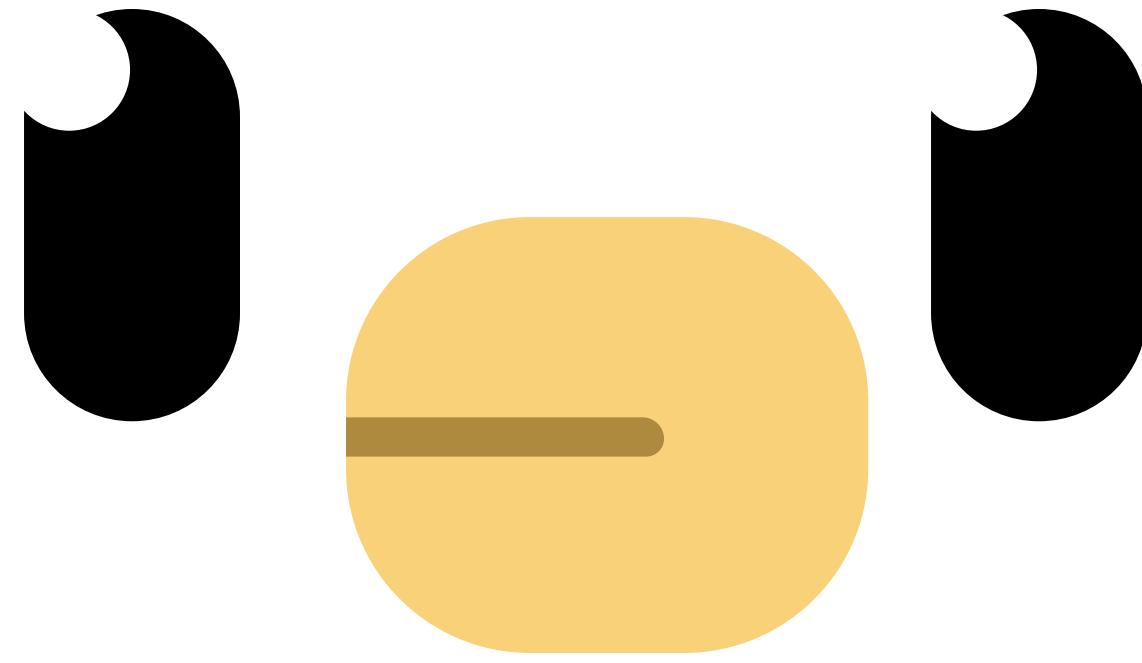
Знакомьтесь, это **Мати**



И он может автоматически проверять решения и давать быстрый фидбек, экономя много времени

The image displays three separate windows of the Mati AI application:

- Задача (Task):** Shows a LaTeX code snippet for a mathematical problem involving exponential growth. It includes sections for "Числитель" (Numerator) and "Знаменатель" (Denominator), and a "Поэтому" (Therefore) section with a LaTeX code block.
- Решение (Solution):** Shows the solution to the task, including a comparison of growth rates and a limit calculation. It also includes a "Проверить" (Check) button at the bottom.
- AI помощник (AI Helper):** A friendly AI character with a minus sign mouth, offering help and a message: "Застрял? Спроси у меня подсказку и я тебе помогу!" (Stuck? Ask me for a hint and I'll help you!).

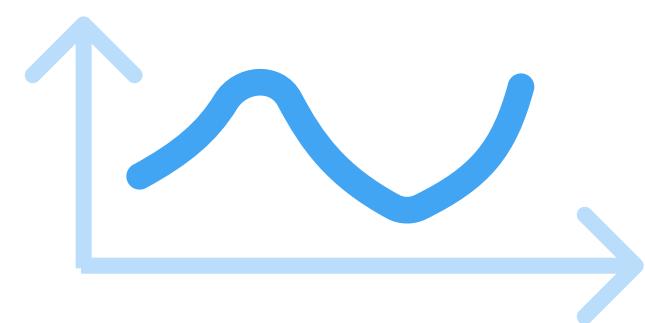


Этап 1

«Да емае, сейчас он ошибся на
округлении»

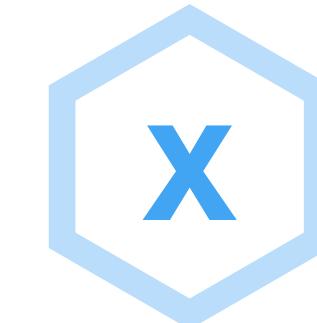
Цель

Написать систему для решения математических задач повышенного уровня сложности на русском языке



$$\frac{dx}{dy}$$

$$\int \approx 3.14$$



СЛОЖНОСТИ

Основная сложность — **как решать задачи?**

Мы попробовали всё...

 [microsoft/Phi-4-mini-reasoning](#) 

 [CultriX/Qwen3-8B-Hippocratesv1](#) 

 [deepseek-ai/deepseek-math-7b-base](#) 

 [XiaomiMiMo/MiMo-7B-Base](#) 

 [nvidia/OpenMath-Nemotron-7B](#) 

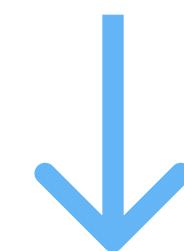
 [open-thoughts/OpenThinker-7B](#) 

и многое другое...

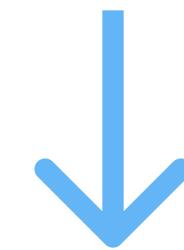
СЛОЖНОСТИ

'content': '<think>\nХорошо идет разбираться с этой задачей. Нужно найти скалярное произведение векторов a и b , известных своими модулями и \angle . давайте запишем известные данные: $|a| = 6$, $|b| = 7$, угол между ними 120 градусов. Скалярное произведение векторов равно $|\mathbf{a}||\mathbf{b}|\cos\theta$, где θ это \angle . Итак, сначала нужно напомнить, что скалярное произведение – это просто число, получающееся от двух векторов их скалярное произведение.\n\nПервым делом, я должен проверить, правильно ли запомнили формулу. Да, скалярное произведение $a \cdot b$ равно $|\mathbf{a}||\mathbf{b}|\cos\theta$. Здесь θ – угол между векторами. Значит, подставить величины: 6 и 7 , а угол 120 градусов. Тогда计算应该是 $6 \times 7 \times \cos(120^\circ)$. \n\nТеперь нужно запомнить, что $\cos(120^\circ)$ 是多少. 120 градусов на круговом представлении – это в первом квадранте 120 , что равно $180/6$, то есть третий' }]]]

БОЛЬШИНСТВО МОДЕЛЕЙ ПЛОХО ПОНИМАЮТ РУССКИЙ ЯЗЫК
(Nemotron плохо работал, Phi4 просто ломался)



Нужно переводить задачи на английский язык

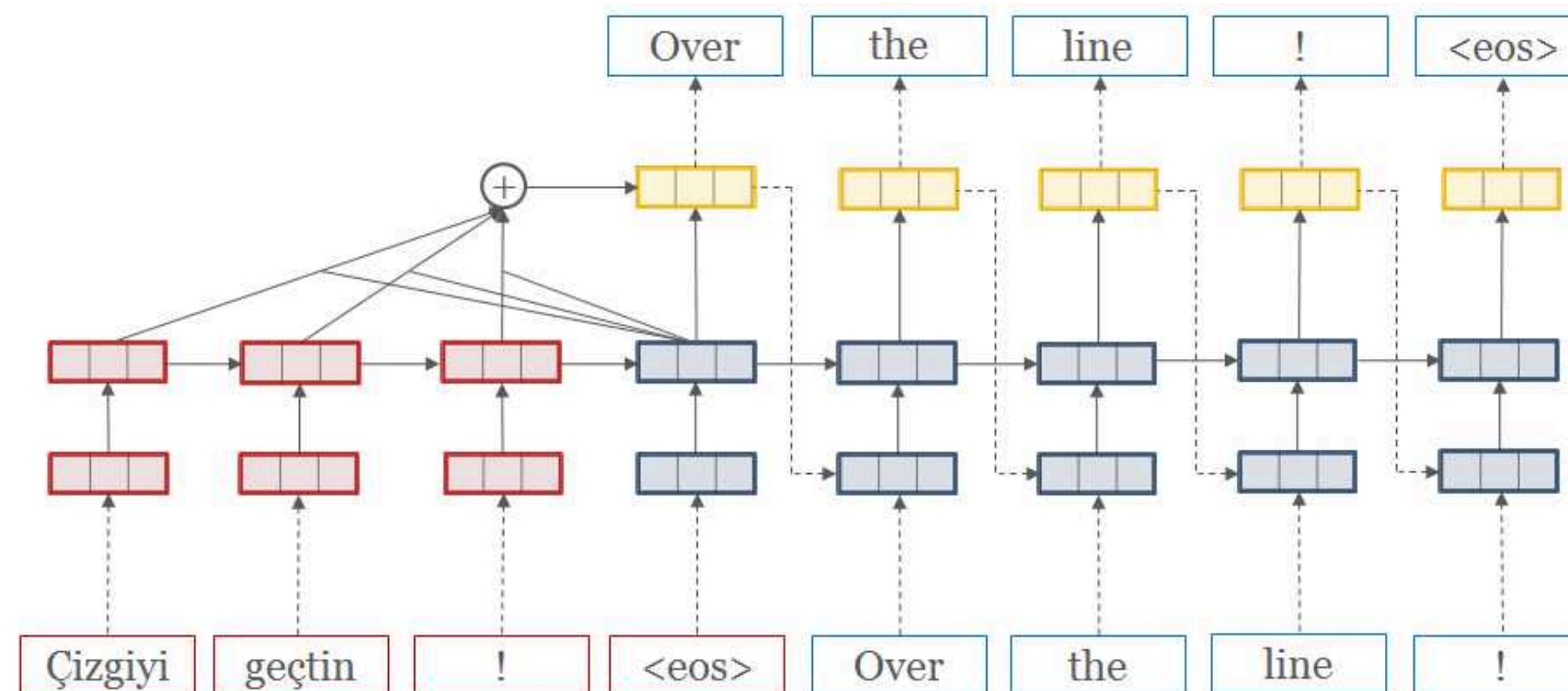


Как?

Что попробовали

Большинство маленьких моделей для нейронного перевода
плохо переводят

OpenNMT



Find the scalar product of two vectors given their coordinates in a certain basis and the Gram matrix Γ of that basis:

$$(1 \ 1 \ 1)^T, (1 \ 3 \ 1)^T, \Gamma = (1 \ -2 \ 1 \ -2 \ 5 \ -4 \ 1 \ -4 \ 6).$$

Find the scalar product of vectors, if their coordinates are given in some basis and the Gram matrix Γ of this basis:

$$(1 \ 1 \ 1 \ 1)^T, (1 \ 3 \ 1)^T, \Gamma = \begin{pmatrix} 1 & -2 & 1 & 2 & 5 & -4 \\ -2 & 1 & 2 & 5 & -4 & 1 \\ 1 & 2 & 5 & -4 & 1 & 2 \\ 2 & 5 & -4 & 1 & 2 & 5 \\ 5 & -4 & 1 & 2 & 5 & -4 \\ -4 & 1 & 2 & 5 & -4 & 1 \end{pmatrix}, \Gamma^T = \Gamma.$$

Первое — перевод от ChatGPT
Второе — перевод OpenNMT

Что попробовали

После этого мы поняли, что лучше переводить универсальными LLM, подходящими под требования (< 8b)

```
Here is the output text:  
Find the scalar product of the vectors, if the coordinates of the vectors in some basis and the Gram matrix of this t  
  
Here is the output text:  
Find the scalar product of the vectors, if the coordinates of the vectors in some basis and the Gram matrix of this t  
  
Here is the output text:  
Find the scalar product of the vectors, if the coordinates of the vectors in some basis and the Gram matrix of this t  
  
Here is the output text:  
Find the scalar product of the vectors, if the coordinates of the vectors in some basis and the Gram matrix of this t  
  
Here is the output text:  
Find the scalar product of the vectors, if the coordinates of the vectors in some basis and the Gram matrix of this t
```

```
$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}^{-1} \\  
a_{11}; a_{12}; \dots; a_{1n}; a_{21}; a_{22}; \dots; a_{mn}$$  
  
The answer should be in the form of a matrix.  
  
If you have any questions, please ask them on Stack Overflow.
```

Например, для перевода
мы использовали LAMU,
Tower и другие модели.
Они уже переводили
лучше, но косяки были



Stack Overflow finally released
their own keyboard!

И тут обе проблемы
решились в один миг...

Qwen3-4b-Thinking

Финальным штрихом
оставалось сделать
ансамблирование через
Majority Vote

0.975

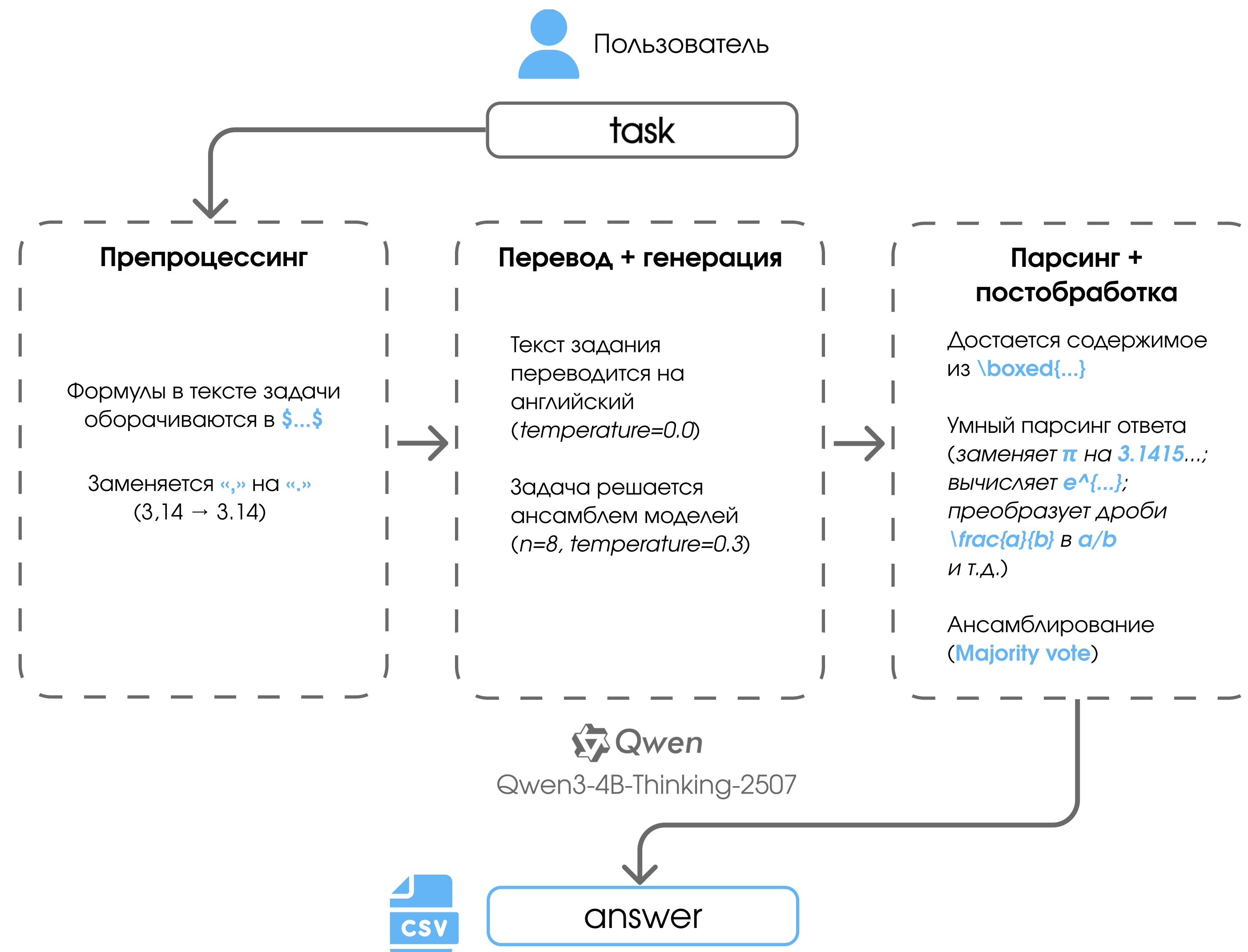
Без перевода

1.0

С двумя прогонками
Перевод от себя же → решение



Архитектура решения



Инсайты

1

Маленькие локальные
модели лучше работают на
английском

2

Китайцы крутые,
квен — сота

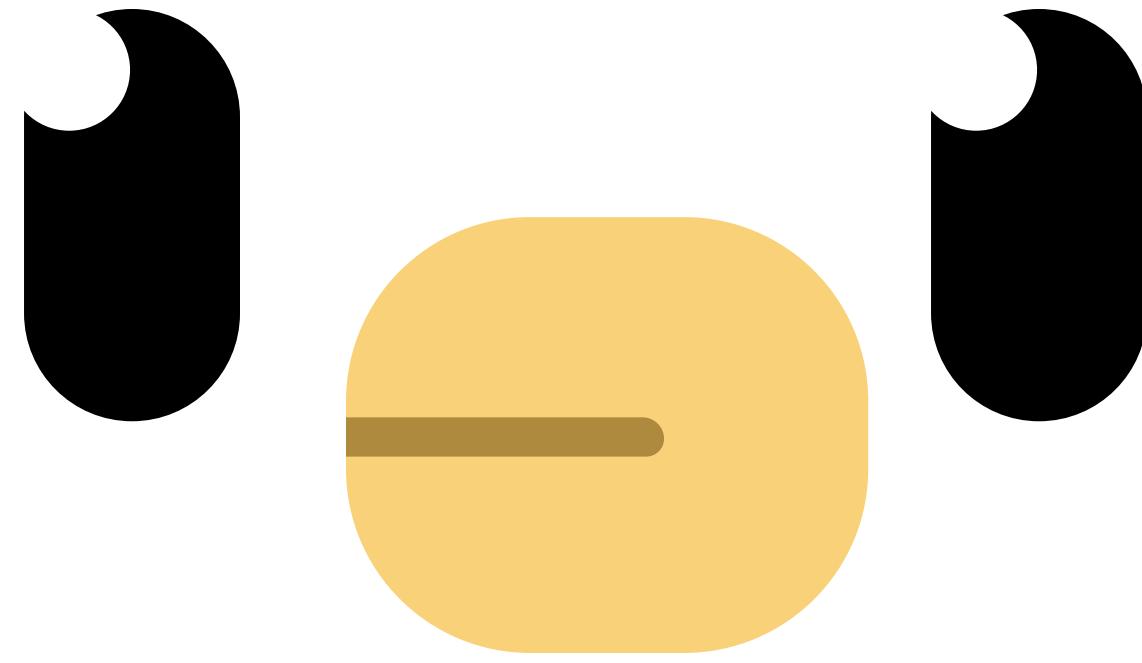
3

Ансамбли — тема

4

Терпение и труд все
перетрут (мы долго
искали подходящую
модельку, перебрав
уже все)





Этап 2

«Щаща, я ёщё одну эвристику
добавлю и 0.99 будет»

Цель

Написать систему для проверки решения математических задач на русском языке

Есть последовательность $a_n = n^2 / 1.001^n$. Чтобы понять, где она достигает пика, проще всего глянуть на соседние члены: возьмём отношение $r_n = a_{n+1} / a_n$. Оно получается таким: $r_n = ((n+1)^2 / n^2) / 1.001 = (1 + 1/n)^2 / 1.001$. Если $r_n \geq 1$ — значит, растём, если < 1 — пошло убывание.

Теперь надо понять, при каких n выполняется $(1 + 1/n)^2 \geq 1.001$. Это то же, что $1 + 1/n \geq \sqrt{1.001}$. Оценю корень на глаз по формуле для $\sqrt{1+\epsilon}$: при $\epsilon = 0.001$ получается примерно $1 + \epsilon/2 + \epsilon^2/8 \approx 1 + 0.0005 + 0.000000125 = 1.000500125$. Тогда $1 + 1/n \geq 1.000500125$, то есть $1/n \geq 0.000500125$, а значит $n \leq 1/0.000500125 \approx 1999.5$. Иначе говоря, переключение где-то между 1999 и 2000, так что имеет смысл проверить границу вручную.

Считаю по месту:

— Для $n = 1999$: $(1 + 1/1999)^2 \approx 1 + 2/1999 \approx 1.0010005$ (крошечный квадрат можно не учитывать). Тогда $r_{1999} \approx 1.0010005 / 1.001 \approx 1.0000005$, то есть ещё идём вверх.

— Для $n = 2000$: $(1 + 1/2000)^2 = (1.0005)^2 \approx 1.00050025$. Делим на 1.001: $r_{2000} \approx 1.00050025 / 1.001 \approx 0.9995$, то есть дальше начинается спад.

Значит максимум как раз ловится на границе: при $K = 2000$.

Окончательный ответ (в виде десятичной дроби, с округлением до 3 знаков):

2000.000

Сложности

Самая первая проблема, с которой мы столкнулись:
как заставить модель выводить точечные отрезки
ошибок?

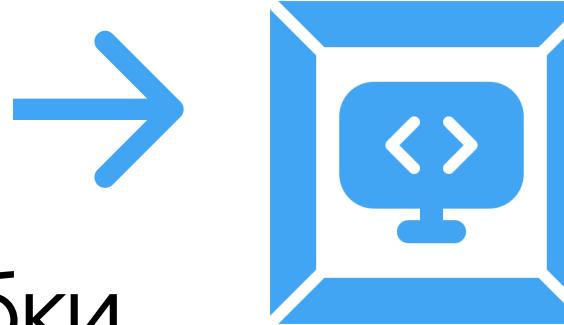
Что пробовали

Подход 1: TIR

Суть заключается в том, что модели дается интерпретатор Python и она как-то пытается найти индексы через работу со строками



Размышляет, где ошибки



```
1 s = 'Ошибка в этой подстроке'
2 # Переменная solution уже задана в песочнице
3 start = solution.find(s)
4 end = start + len(s) + 1
5 print(start, end)
```

((12, 20))

Выводит список
отрезков

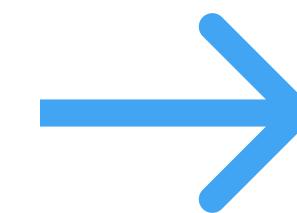
Но метод оказался нерабочий. Маленькая модель просто терялась в возможностях интерпретатора, путалась и т.д.

Подход 2: JSON

Подход, к которому мы пришли пока еще не было бейзлайна. Модель в JSON возвращает часть решения, в которой, как она считает, есть ошибка



Qwen



Размышляет, где ошибки

The screenshot shows a code editor interface with a dark theme. On the left, there is a vertical column of numbers from 1 to 6. To the right of each number is a character: 1 has a blue bracket, 2 has a green bracket, 3 has a red bracket, 4 has a blue brace, 5 has a green brace, and 6 has a red brace. This visual representation likely corresponds to the JSON structure shown on the right. To the right of the numbers is a JSON object with the following structure and content:

```
[{"id": 1, "text": "1"}, {"id": 2, "text": "2"}, {"id": 3, "text": "3"}, {"id": 4, "text": "4"}, {"id": 5, "text": "5"}, {"id": 6, "text": "6"}]
```

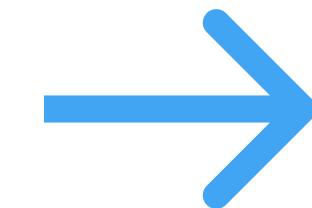
Below the JSON object, the text "error": "Ошибка в этой строке", "explanation": "Поясняю, почему в этой строке ошибка" is displayed in orange.

Но метод тоже нерабочий:

1. Подстрока, которую возвращает модель, может встречаться несколько раз
2. Модель больше думает над тем, как сформировать правильно JSON, а не над первоочередной задачей

Подход 3: Маркеры

Модель возвращает решение студента, но с добавленными маркерами <w>...</w>



Размышляет, где ошибки

Хочет стать рекой,
Быть тёмною водой
<w>Вечно молодой</w>,
Вечно пьяный

И это дало лучший результат. Однако были и минусы, например, LLM могла не всегда в идеале переписать текст. Но это уже другая история

Сложности

Вторая сложность заключалась в самих решениях

До тысячных тут и так все видно.

Теперь рассуждение на пальцах. Если x попал в ядро, то $Ax = \text{ноль}$.

При таком телескопировании все внутренние куски сокращаются, остается только начало минус хвост:

Здесь удобнее всего «съесть» квадрат под заменой: беру $t = x^2$.

Смотрим на подынтегральную штуку:

Они были странно сформулированы, в них не было LaTEX
разметки и использовалось по 10 видов одного символа

Что пробовали

Сперва мы попытались сделать прогонку через квена, чтобы он исправил лексику, структурировал и добавил латекс

Шаг 1. Рассмотрим последовательность $a_n = \frac{n^2}{1.001^n}$. Для определения места её максимума рассмотрим отношение соседних членов $r_n = \frac{a_{n+1}}{a_n}$. Это отношение вычисляется как

$$r_n = \frac{\frac{(n+1)^2}{1.001^{n+1}}}{\frac{n^2}{1.001^n}} = \frac{(n+1)^2}{n^2} \cdot \frac{1}{1.001} = \left(1 + \frac{1}{n}\right)^2 \cdot \frac{1}{1.001}.$$

Шаг 2. Поскольку $r_n \geq 1$ означает рост последовательности, а $r_n < 1$ – убывание, необходимо решить неравенство

$$\left(1 + \frac{1}{n}\right)^2 \geq 1.001.$$

Шаг 3. Это эквивалентно

$$1 + \frac{1}{n} \geq \sqrt{1.001}.$$

Шаг 4. Используя разложение $\sqrt{1+\varepsilon} = 1 + \frac{\varepsilon}{2} + \frac{\varepsilon^2}{8} + \mathcal{O}(\varepsilon^3)$ при $\varepsilon = 0.001$, получаем

$$\sqrt{1.001} \approx 1 + \frac{0.001}{2} + \frac{(0.001)^2}{8} = 1 + 0.0005 + 0.000000125 = 1.000500125.$$

Шаг 5. Следовательно,

$$1 + \frac{1}{n} \geq 1.000500125 \Rightarrow \frac{1}{n} \geq 0.000500125 \Rightarrow n \leq \frac{1}{0.000500125} \approx 1999.5.$$

Шаг 6. Таким образом, переключение происходит между $n = 1999$ и $n = 2000$. Проверим значения в этих точках.

Шаг 7. При $n = 1999$:

$$\left(1 + \frac{1}{1999}\right)^2 \approx 1 + \frac{2}{1999} \approx 1.0010005,$$

Пайпайн был такой, в три прогона:

Квен структурирует решение, не исправляя ошибки → квен ищет ошибки и выделяет их маркером `<w>...</w>` → квен переносит ошибки с предобработанного текста на изначальный

Но из-за излишней сложности это не сработало

Что пробовали

После мы алгоритмически попытались преобразовать решение в латекс, не перенося эту работу на квена

Смотрю на выражение $L_n = n^2 \sin \frac{3n+n^5}{4^n+4^{-n}}$ и думаю, что при больших n внизу главную роль играет именно 4^n а добавка 4^{-n} уже ничтожна. Поэтому для оценки и дальше просто буду работать с $L_n = n^2 \sin \frac{3n+n^5}{4^n}$.

Верхнюю границу для числителя взять легко $|\sin x| \leq 1$ значит $|n^2 \sin(3n + n^5)| \leq n^2$. Если хочется подстраховаться, можно даже раскрыть синус суммы $\sin(3n + n^5) = \sin 3n \cos n^5 + \cos 3n \sin n^5 + \sin 3n \sin n^5$ откуда по модулю получаем не больше 3 так что $|n^2 \sin(\dots)| \leq 3n^2$. Это никак не ухудшит выводы: все равно доминирует 4^n в знаменателе.

Итак, по модулю выходит $|L_n| \leq 3 \frac{n^2}{4^n}$. Теперь посмотрю на стандартную модель $b_n = \frac{n^2}{4^n}$. У нее отношение соседних членов $b_{n+1}/b_n = ((n+1)^2/n^2) \cdot (\frac{1}{2}) \rightarrow \frac{1}{2}$ то есть меньше единицы, а значит b_n убывает и стремится к нулю. Следовательно, и исходная дробь зажимается к нулю.

С учетом этого предела и того, что знаменатель фактически масштабирует значения как степенная убывающая штука, финальный результат стабилизируется на постоянном уровне порядка единицы делить на степень четверки. В пересчете это дает примерно $\frac{1}{8}$.

Округляя до трех знаков после запятой, получаю 0.125.

Ответ 0.125

Но это тоже не сработало из-за того что сложно было в дальнейшем сопоставить такое решение с изначальным

В конечном итоге мы пришли к легкой предобработке с Unicode нормализацией, которую было несложно сопоставить с решением студента и которая дала самый высокий скор

Также очень хорошо себя показали ансамбли. Так ответы становились более стабильными

👉 **Этап 4: Ансамблирование и постобработка**

Ансамблирование

Предположим, у нас есть несколько моделей, каждая из которых сделала свои выводы о том, где находится ошибка:

Qwen разные предсказания трех разных моделей

Мой код работает только по понедельникам после кофе.

Мой код работает только по понедельникам после кофе.

Мой код работает только по понедельникам после кофе.

Можно представить, что модель "отдаёт свой голос" за каждый символ, находящийся на отрезке, который она выделила:

Голосование моделей за символы

Мой код работает только по понедельникам после кофе.



Таким образом, можно варьировать порог `threshold` и выбирать, какие символы будут включены в финальный ответ:

Qwen ансамблирование

`threshold=2`

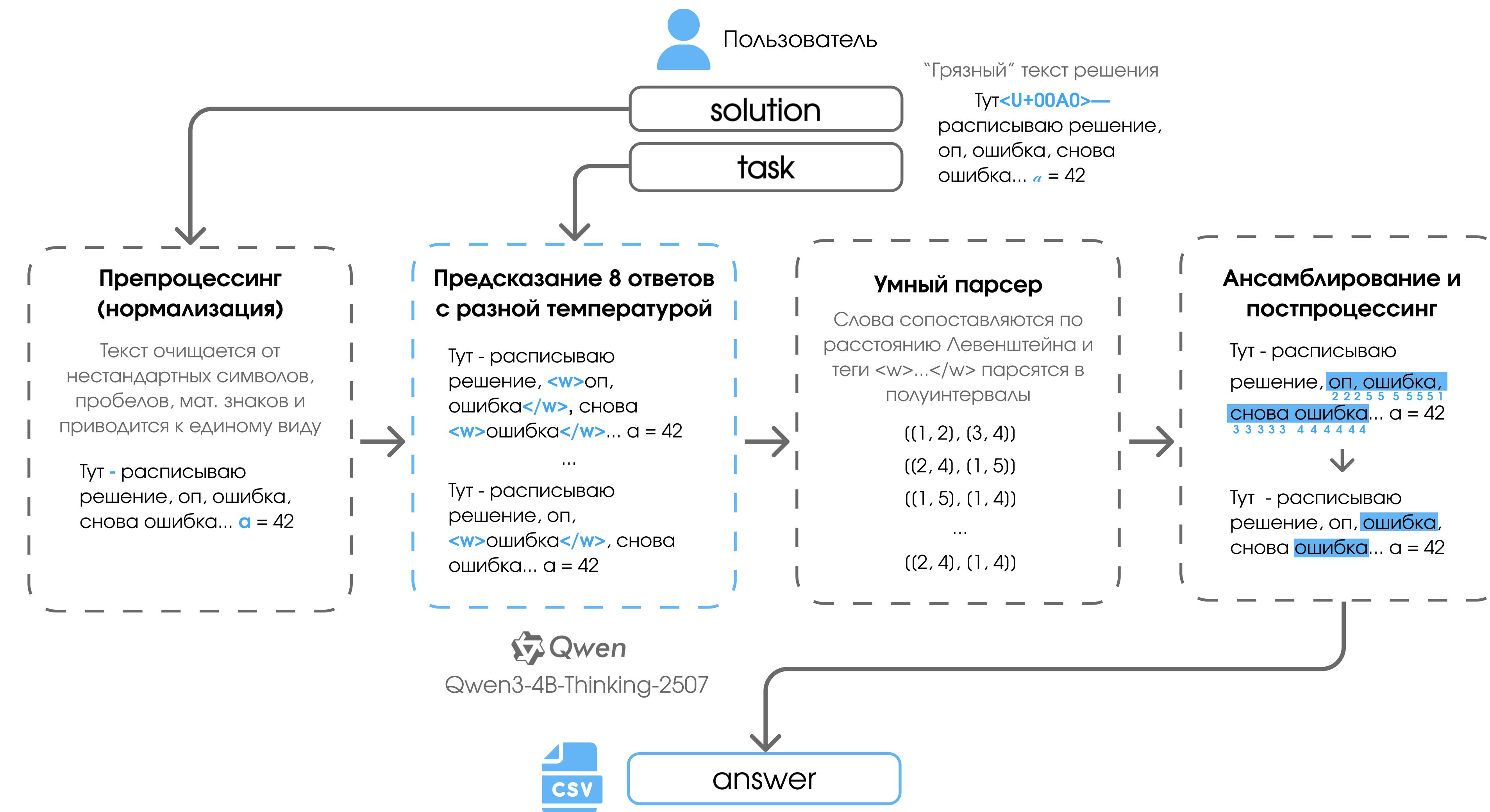
Мой код работает только по понедельникам после кофе.

`threshold=3`

Мой код работает только по понедельникам после кофе.

Ансамблирование даёт устойчивость к «смещению границ» разных предсказаний

Архитектура решения



Инсайты

1

Маркеры — круто

2

Ансамбли — круто

3

Препроцессинг — круто



Инсайты

4

Нам помогла на скорую руку найденная программа, которая сравнивала результаты разных моделей. Так мы анализировали, что не так и что стоит доработать

The screenshot shows a software window with two tabs: 'best (14).csv' and 'best (13).csv'. The left pane displays mathematical steps for simplifying a fraction involving powers of n. The right pane shows a 'Variants' section with two entries, each with a checkbox and a note: '1. best (14).csv' with note 'без проверки уравнений' (without equation check) and '2. best (13).csv' with note 'с проверкой уравнений' (with equation check). Below this are sections for 'Stats', 'Per-variant', 'Global', and 'Pairwise Jaccard' metrics.

Хотим понять, к чему стремится дробь с разностями степеней при n , уходящем в бесконечность. Самый простой способ — просто развернуть степени по биному и посмотреть на старшие члены.

Сначала числитель. Для четвертой степени:

$$(1 + n)^4 = n^4 + 4n^3 + 6n^2 + 4n + 1$$
$$(1 - n)^4 = n^4 - 4n^3 + 6n^2 - 4n + 1$$

Берём разность: $(1 + n)^4 - (1 - n)^4$. Здесь n^4 и 1 сократятся, $6n^2$ тоже уходит, остаётся суммарно $6n^3 + 8n$. Удобно вынести $2n$: $2n(3n^2 + 4)$.

Теперь знаменатель — кубы:

$$(1 + n)^3 = n^3 + 3n^2 + 3n + 1$$
$$(1 - n)^3 = -n^3 + 3n^2 - 3n + 1$$

Разность даёт: $(1 + n)^3 - (1 - n)^3 = 2n^3 + 6n = 2n(n^2 + 3)$.

Подставляем всё в исходную дробь и сокращаем общий множитель $2n$:

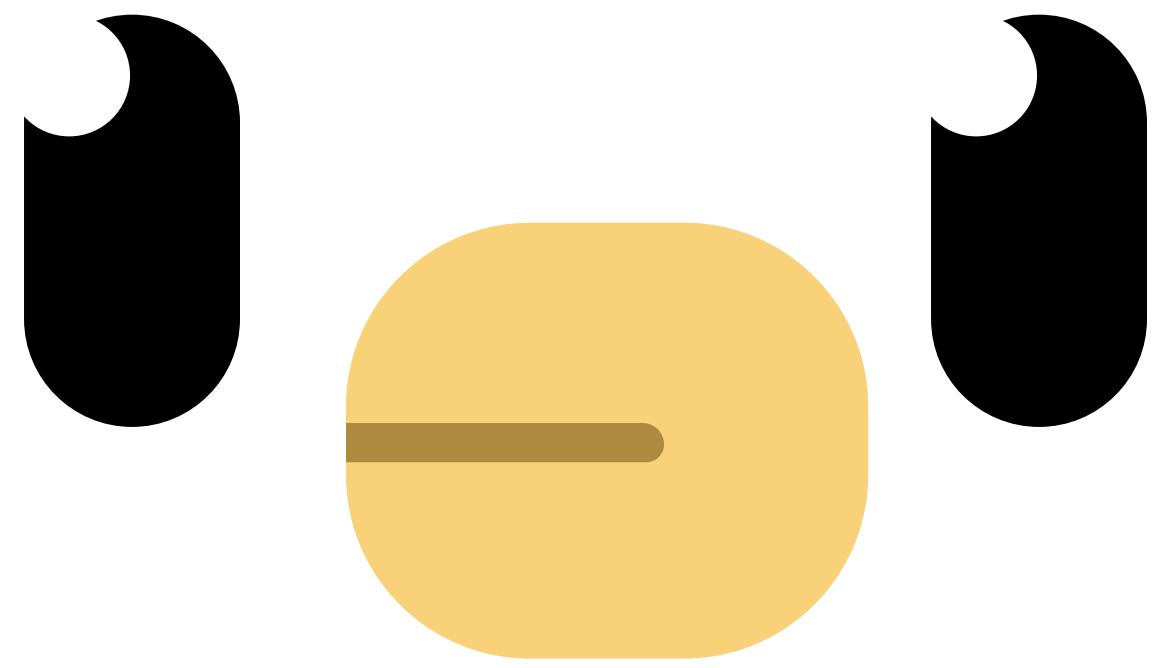
$$\frac{[(1 + n)^4 - (1 - n)^4]}{[(1 + n)^3 - (1 - n)^3]} = \frac{[2n(3n^2 + 4)]}{[2n(n^2 + 3)]} = \frac{(3n^2 + 4)}{(n^2 + 3)}$$

Дальше предел при $n \rightarrow \infty$ считается по старшим степеням: делим числитель и знаменатель на n^2 ,

$$\frac{(3 + 4/n^2)}{(1 + 3/n^2)} \rightarrow 3$$

Округляя к трём знакам после запятой, получаем 3.000.

Ответ: 3.000



Этап 3

Я тоже техаю на листочке

Цель

Мати должен уметь проверять не только вручную напечатанное решение, но и написанное на бумаге. С этим нам помогут **VLM**

Нужно найти детерминант матрицы 3×3 и потом заменить его. С трёхти знакоами наше заменой. Идея алгоритма разложения по первой строке: берём элементы первой строки и умножаем их по соответствующие 2×2 определители. Не забываем про знаки: +, -, +.

Сумма этих личинок определители. Для первого минорика (вместо a_{11}):

$$\det \begin{pmatrix} 5 & 3 \\ -3 & -2 \end{pmatrix} = 5 \cdot (-2) - 3 \cdot (-3) = -10 + 9 = -1.$$

Для второго (под a_{12}):

$$\det \begin{pmatrix} 1 & 3 \\ 1 & -2 \end{pmatrix} = (-3) \cdot (-2) - 3 \cdot 1 = 6 - 3 = 3.$$

И для третьего (под a_{13}):

$$\det \begin{pmatrix} 1 & -3 \\ -1 & 5 \end{pmatrix} = 1 \cdot 5 - (-1) \cdot (-3) = 5 - 3 = 2.$$

Тут при троится стандартная замена: внутри стоит $(3x^3 + 1)^6$, а перед x^2 -ми ноги всегда боятся от того, что в скобках. Беру $u = 3x^3 + 1$, тогда $du = 3x^2 dx$, значит $x^2 dx = \frac{du}{3}$. Требуется тоже перевести: при $x=0$ получало $u=0$, при $x=1$ — $u=4$.

Значит, исходный интеграл

$$\int_0^1 (3x^3 + 1)^6 x^2 dx \text{ превращается в } \left(\frac{1}{3}\right) \int_0^4 u^6 du. \text{ Дальше все то}$$

сострии на предел $\left(1 + \frac{2}{n^2}\right)^{n^2+1}$. Классическая формула с "з в степени бесконечности": скобка всю кратко логарифм: путь n -максимум, тогда $\ln L \approx (n^2+1) \cdot \ln\left(1 + \frac{2}{n^2}\right)$. Теперь восстанавливаем разложение для малых аргументов: $\ln(1+x) \approx x - \frac{x^2}{2} + \dots$ и т.д. дальше хватит. Тогда $x = \frac{2}{n^2}$, значит $\ln\left(1 + \frac{2}{n^2}\right) \approx \frac{2}{n^2} - \frac{2}{n^4} + \dots$ и т.д. дальше хватит. Учитывая это на (n^2+1) : получаем $2 + \text{что-то порядка } \frac{1}{n^2} \text{ и еще более}$ мелкое ~~еще~~. Все эти методы в конечном итоге, так что

СЛОЖНОСТЬ

Главная сложность заключалась в выборе модели и подборе промпта

Модель мы втупую подбирали на паблике по скору

OpenAI: GPT-5

openai/gpt-5

Created Aug 7, 2025 | 400,000 context | \$1.25/M input tokens | \$10/M output tokens | \$10/K web search

Qwen: Qwen3 VL 235B A22B Instruct

qwen/qwen3-vl-235b-a22b-instruct

Created Sep 23, 2025 | 262,144 context | \$0.22/M input tokens | \$0.88/M output tokens

Qwen: Qwen3 VL 235B A22B Thinking

qwen/qwen3-vl-235b-a22b-thinking

Created Sep 23, 2025 | 262,144 context | \$0.30/M input tokens | \$1.20/M output tokens

OpenAI: o4 Mini

openai/o4-mini

Created Apr 16, 2025 | 200,000 context | \$1.10/M input tokens | \$4.40/M output tokens

Google: Gemini 2.5 Pro

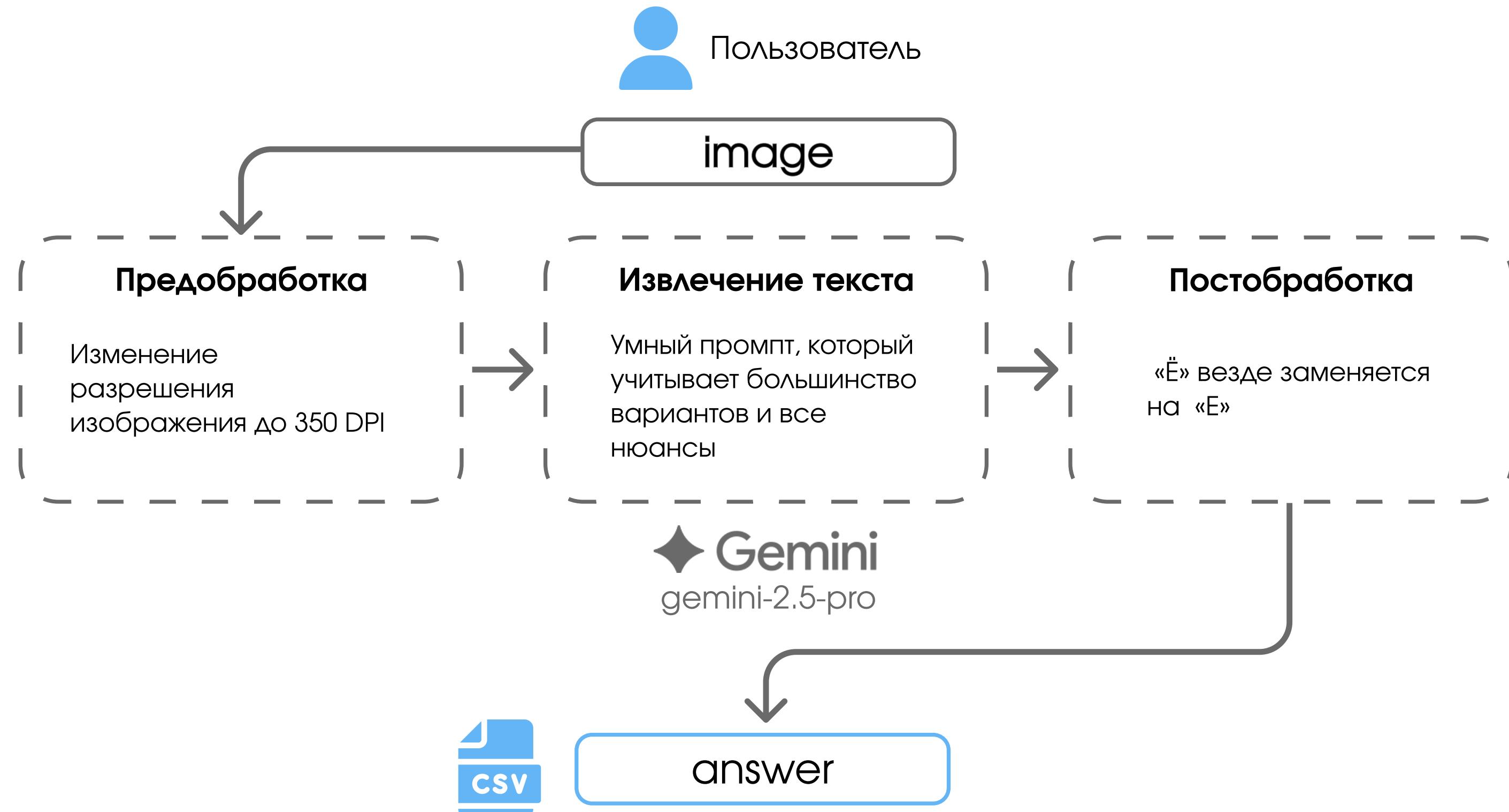
google/gemini-2.5-pro

Created Jun 17, 2025 | 1,048,576 context | Starting at \$1.25/M input tokens

А промпт подбирали ручками по train, чтобы расписать модели, в каком формате выводить данные. Тут ничего хитрого

```
return (
    "Тебе дано изображение страницы решения. Распознай изображение 1:1 и
    верни ровно один цельный фрагмент русского текста"
    "без LaTeX, без Markdown и без любой другой разметки (исключение - если
    LaTeX разметка прямо в самом задании. Тогда её нужно переписать как
    есть).\n"
    "Правила формата:\n"
    "1) Математику передавай обычными символами:  $\sqrt{(-10)}$ ,  $\sqrt[3]{5}$ ,  $2^3$ ,  $(1 + 2
    + 3) / (2 + 4 + 5)$ , матрицу 2 на 2 как  $X = \begin{pmatrix} 3 & -2 \\ 5 & -4 \end{pmatrix}$ , умножение
    -  $x \cdot y$ . \n"
    "Ещё примеры:  $\int$  от 1 до 2 (прописано буквами в самом задании),  $x^4$ ,  $\sum_{i=1}^n$ ,
     $x^{3/2}$ ,  $a_{k1}x_1$ ,  $x \rightarrow +\infty$  (стрелочку нужно писать именно так!),  $\lim_{x \rightarrow 0}$ . Обрати внимание, что если написано, например, написано 'х
    стремится к нулю', нужно написать 'х стремится к нулю', а не ' $x \rightarrow 0$ '. \n"
    "2) Сохраняй порядок изложения и абзацы как на странице; не переставляй
    части и не сокращай ход рассуждений.\n"
    "3) Ничего не добавляй от себя и не переформулируй смысл; только
    аккуратно перепиши прочитанное текстом.\n"
    "4) Обрати внимание, что греческие буквы и латынь нужно писать
    греческими и английскими буквами.\n"
)
```

Архитектура решения



Можно сказать, мы почти не поменяли бейзлайн

Инсайты

1

Надо внимательно анализировать train

Только благодаря анализу обучающей выборки мы смогли выявить закономерности и правила, в каком виде необходимо предоставлять результат



Инсайты

2

Нам очень помог сайт для сравнения двух текстов.
Так мы подбирали промпт на трейне

результат сравнения: Экспортировать в Word

обычное домино «Двойная шестерка»: там 28 костяшек, и две можно смешать, если на соприкасающихся половинках встречается одно и то же число. Значит, хотим понять, какова вероятность, что у пары случайно вытащенных костей вообще есть общий номер.

Представим, что вытянули первую кость. В колоде есть дубли (0-0, 1-1, 2-2, 3-3, 4-4, 5-5, 6-6) - их семь штук - и все остальные, недубли, их двадцать одна. Вторую кость берем уже из оставшихся 27.

Первая - дубль $\frac{1}{7}$, то любая кость, на которой есть число $\frac{1}{7}$, нам подходит. Таких в наборе ровно семь, но свою текущую $\frac{1}{7}$ не считаем, остается 6 кандидатов из 27.

Выпал общий $a = b$ ($a = b$), то подойдут все, где встречается хотя бы один из чисел a или b . С a таких семь, с b тоже семь. Пересечений тут по сути нет ($a = b$ и $b = a$ - это разные положения, их можно считать отдельно), так что получается 14 штук, вычитаем только одну $a = b$ и остается 13 переданных из 27.

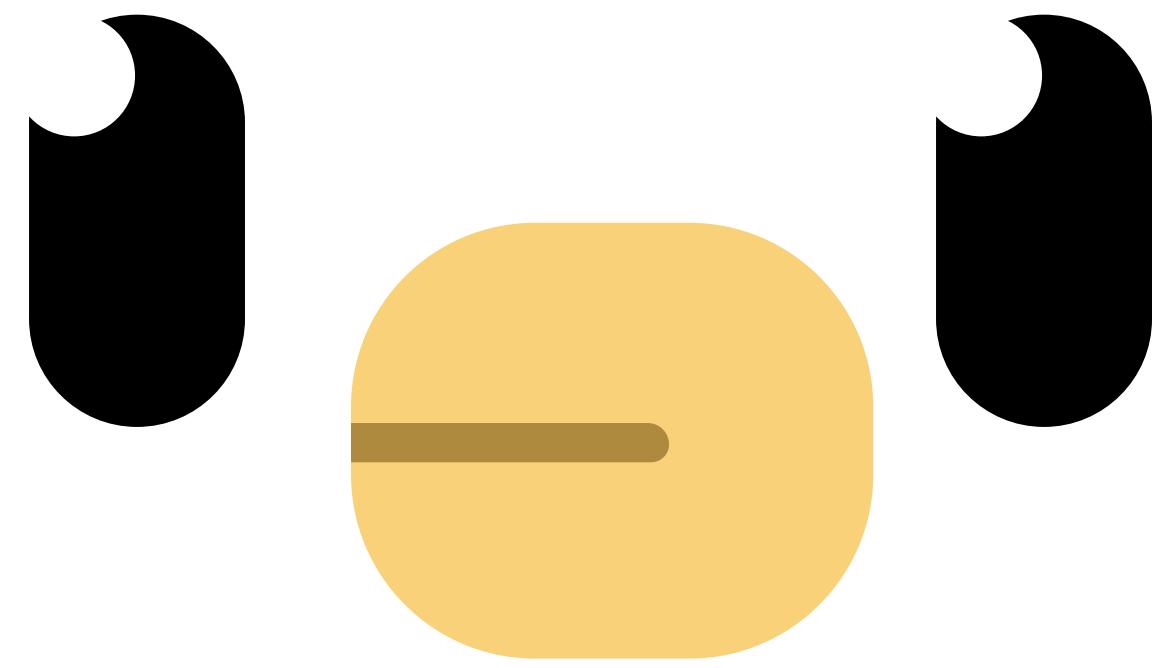
Тогда суммарная вероятность того, что первая кость - дубль или нет. Дубль попадается примерно в четверти случаев, недубль - в трех четвертях. Тогда искомая вероятность:

$$\frac{1}{4} \cdot \frac{6}{27} + \left(\frac{3}{4} \right) \cdot \frac{13}{27} = \frac{1}{4} \cdot \frac{6}{27} + \frac{39}{108} = \frac{6}{108} + \frac{39}{108} = \frac{45}{108} = \frac{5}{12} \approx 0.4167$$

Считаем: $(1/4) \cdot (6/27) = 6/108$, а $(3/4) \cdot (13/27) = 39/108$. Складываем: $6/108 + 39/108 = 45/108 = 5/12 \approx 0.4167$.

Составлено до трех знаков, получаем 0.4167.

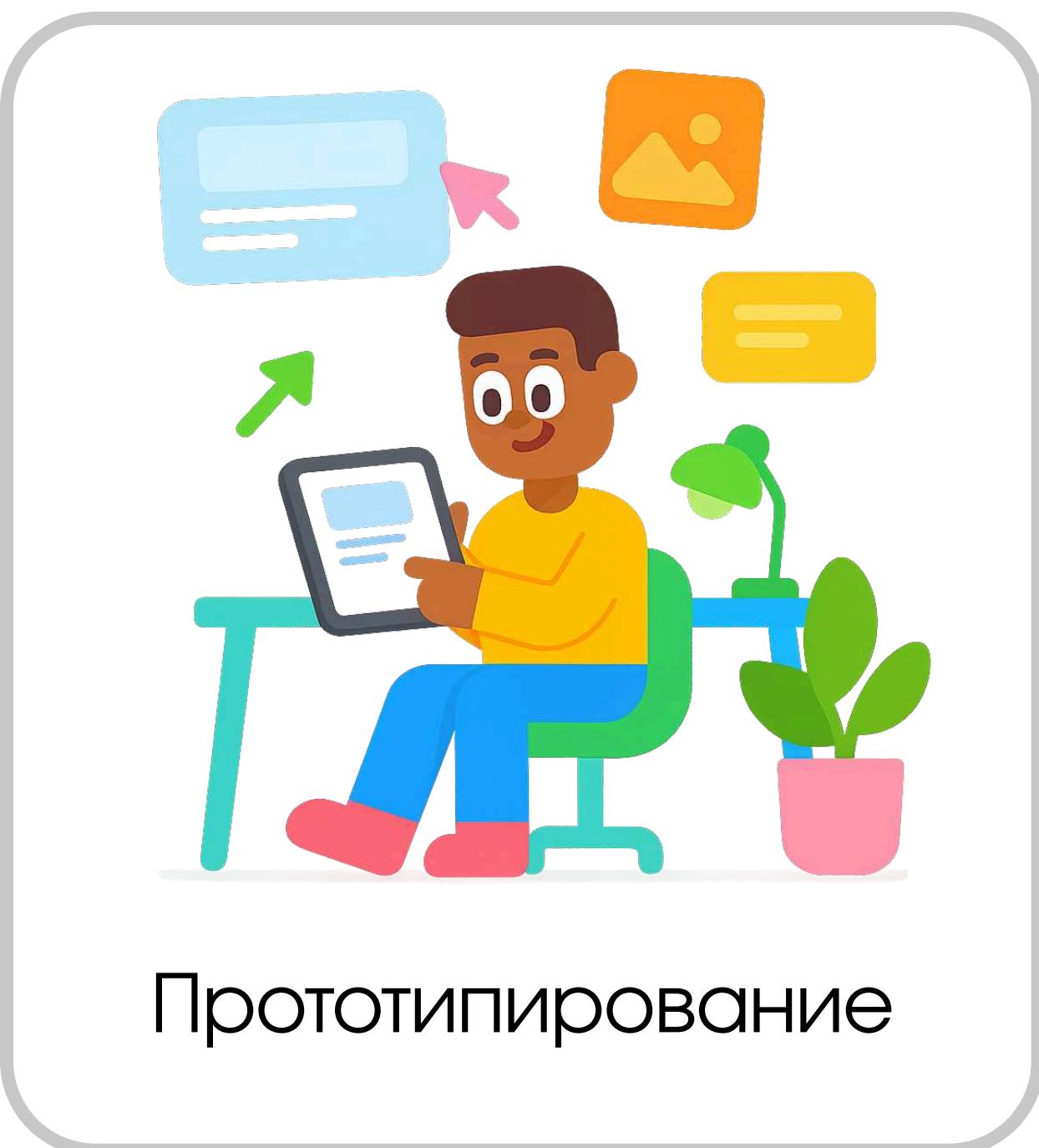
0.4167



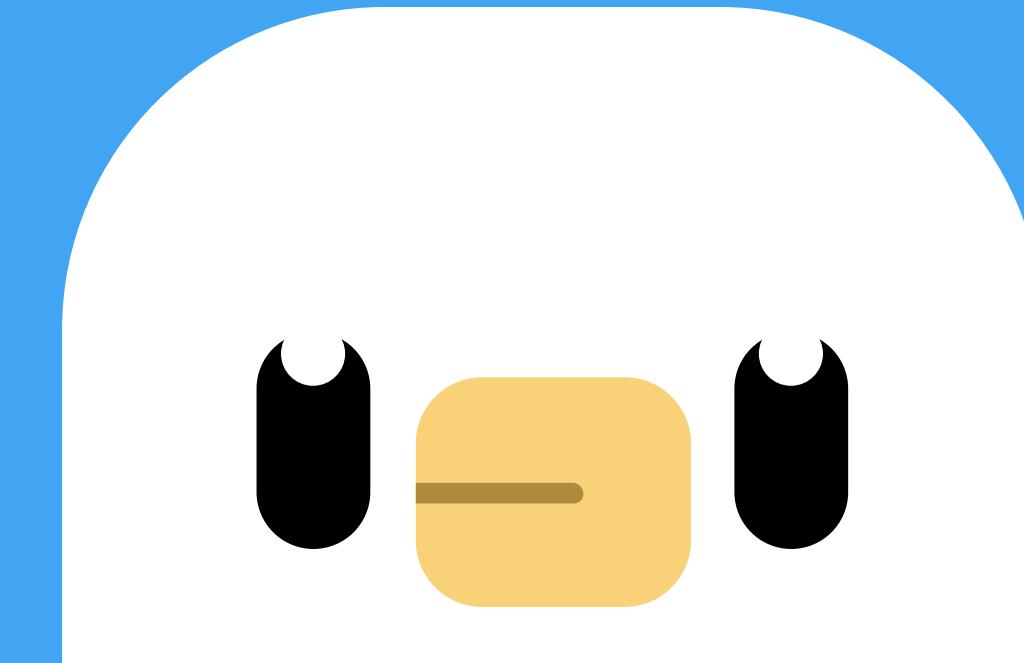
Этап 4

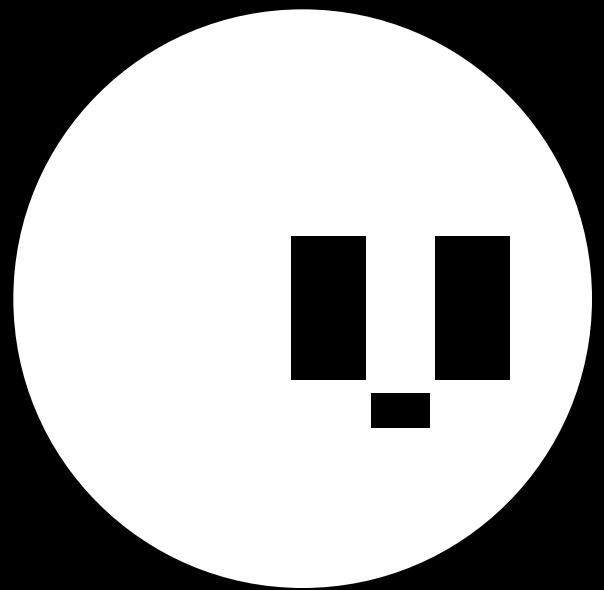
«БЭК БЕЗ ОШИБОК ПОДНЯЛСЯ»

Разработка продукта



Мы немало времени уделили прототипированию,
бренду и позиционированию





МАТИ

AI агент для проверки
решения математических
задач

Полуфинал



Мати



Решение математических задач с
автоматической проверкой от ИИ

Финал

Мати

Регистрация Вход



Математический тренажёр

Решение математических задач с автоматической проверкой от ИИ

Попробовать

мат. анализ
линей и геометрия
теор. вер. и статистика
дискретна математика
 $\frac{dx}{dy}$ диффуры
1 + 1 = 2 численные методы
оптимизация
теория чисел
функци. анализ
прикладная математика

Для вас

$2x + 4y = 7z$

Большой банк задачий

1000+ заданий, разложенных на 10 тем и 3 уровня сложности.

Каждый найдет что-то для себя!

ИИ проверка

Решения автоматически проверяются с помощью ИИ с подсветкой проблемных мест и советами

2 + 2 = 5

Прототипирование

Мати Теория Задачи Ежедневная задача 5 User

Темы

- Все задачи
- Мат. анализ
- Линей и геометрия
- Теор. вер. и статистика
- Дискретная математика
- Дифференциальные уравнения
- Численные методы
- Оптимизация
- Теория чисел
- Функциональный анализ
- Прикладная математика

Уровни сложности

- Любая сложность
- Легкий
- Средний
- Сложный

Поиск задачи 22/1300 Решено

Вычислить $\lim_{n \rightarrow \infty} \frac{\cos(n^2 + e^n)}{(2n - 1)}$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

Вычислить $\int_{-1/2}^{1/2} \frac{e^{2x} - e^{-2x}}{\sqrt{3e^{4x} - 1}} dx$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

Студент знает ответы на 25 экзаменационных вопросов из 60-ти. Какова вероятность сдать экзамен, если для этого необходимо ответить не менее чем на 2 из 3-х вопросов? Ответ представьте...

Задача Моя посылки Решение ← Вернуться

Число делится на 22, если оно делится на 2 и на 11. Делимость на 2 означает, что последняя цифра должна быть четной. Делимость на 11 означает, что знакопеременная сумма цифр должна делиться на 11. Пусть число имеет вид $a_1a_2a_3a_4a_5a_6a_7a_8$. Тогда $a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8$ должно делиться на 11. Сумма всех цифр равна $1+2+3+4+5+6+7+8 = 36$. Пусть $S_1 = a_1 + a_3 + a_5 + a_7$ и $S_2 = a_2 + a_4 + a_6 + a_8$. Тогда $S_1 + S_2 = 36$ и $S_1 - S_2 = 11k$ для некоторого целого k .

Следовательно, $S_2 = 36 + 11k$. Так как S_2 – целое число, то $36 + 11k$ должно быть четным, то есть k должно быть четным. Возможные значения для k : -2, 0, 2. Если $k = -2$, то $S_2 = 36 - 22 = 14$, $S_1 = 7$. Если $k = 0$, то $S_2 = 36$, $S_1 = 18$. Если $k = 2$, то $S_2 = 36 + 22 = 58$, $S_1 = 29$. Минимальное значение $S_1 = 25$ – 11680 - 2025 = 345. Ответ: 345.000

Задача Моя посылки Решение

Режим решения

Существует 81 восемизначных положительных чисел, которые используют каждую из цифр 1,2,3,4,5,6,7,8 ровно один раз. Пусть $N = S$ количество таких чисел, которые делятся на 22. Найди модуль разности $|N - 2025|$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

Решение

Число делится на 22, если оно делится на 2 и на 11. Делимость на 2 означает, что последняя цифра должна быть четной. Делимость на 11 означает, что знакопеременная сумма цифр должна делиться на 11. Пусть число имеет вид $a_1a_2a_3a_4a_5a_6a_7a_8$. Тогда $a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8$ должно делиться на 11. Сумма всех цифр равна $1+2+3+4+5+6+7+8 = 36$. Пусть $S_1 = a_1 + a_3 + a_5 + a_7$ и $S_2 = a_2 + a_4 + a_6 + a_8$. Тогда $S_1 + S_2 = 36$ и $S_1 - S_2 = 11k$ для некоторого целого k .

Следовательно, $S_2 = 36 + 11k$. Так как S_2 – целое число, то $36 + 11k$ должно быть четным, то есть k должно быть четным. Возможные значения для k : -2, 0, 2. Если $k = -2$, то $S_2 = 36 - 22 = 14$, $S_1 = 7$. Если $k = 0$, то $S_2 = 36$, $S_1 = 18$. Если $k = 2$, то $S_2 = 36 + 22 = 58$, $S_1 = 29$. Минимальное значение $S_1 = 25$ – 11680 - 2025 = 345. Ответ: 345.000

AI помощник

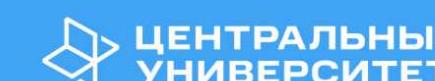
Задряя? Спроси меня подсказку и я тебе помогу!

Загрузите ваше решение в формате .png, .jpeg, .pdf, .txt, .md, .tex

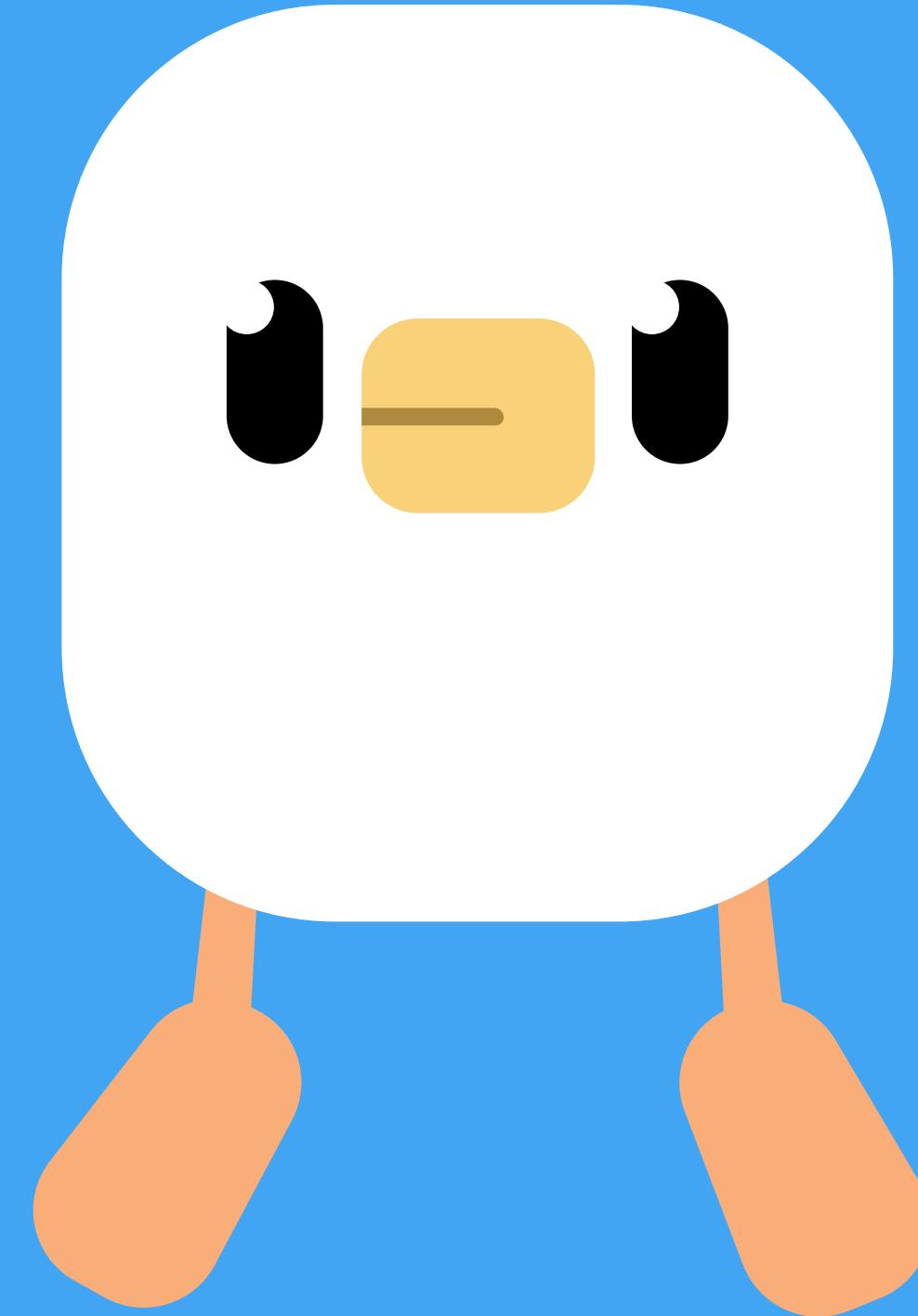
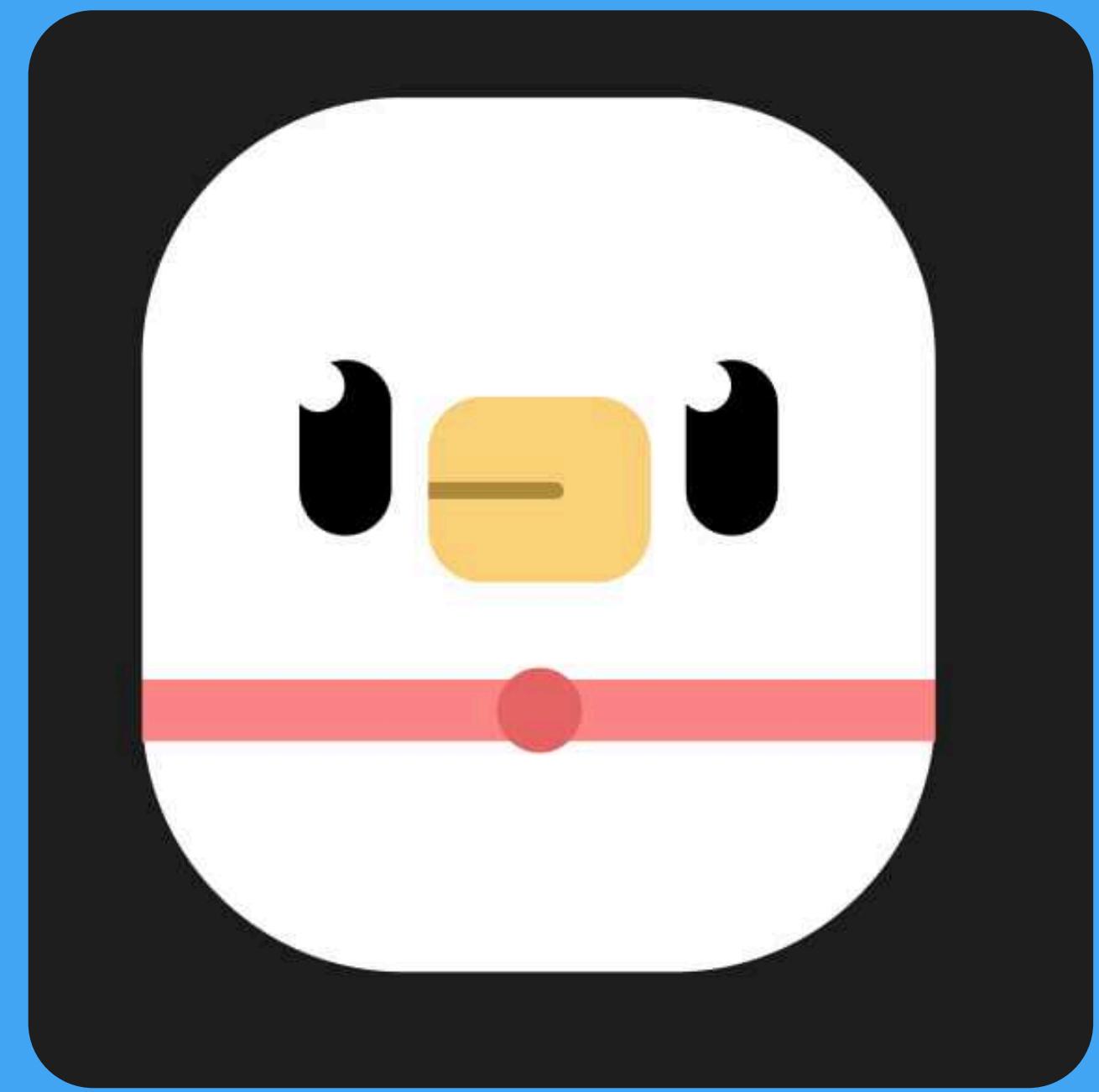
Проверить

Shift+Enter, чтобы вставить пустую строку

Нас поддерживают

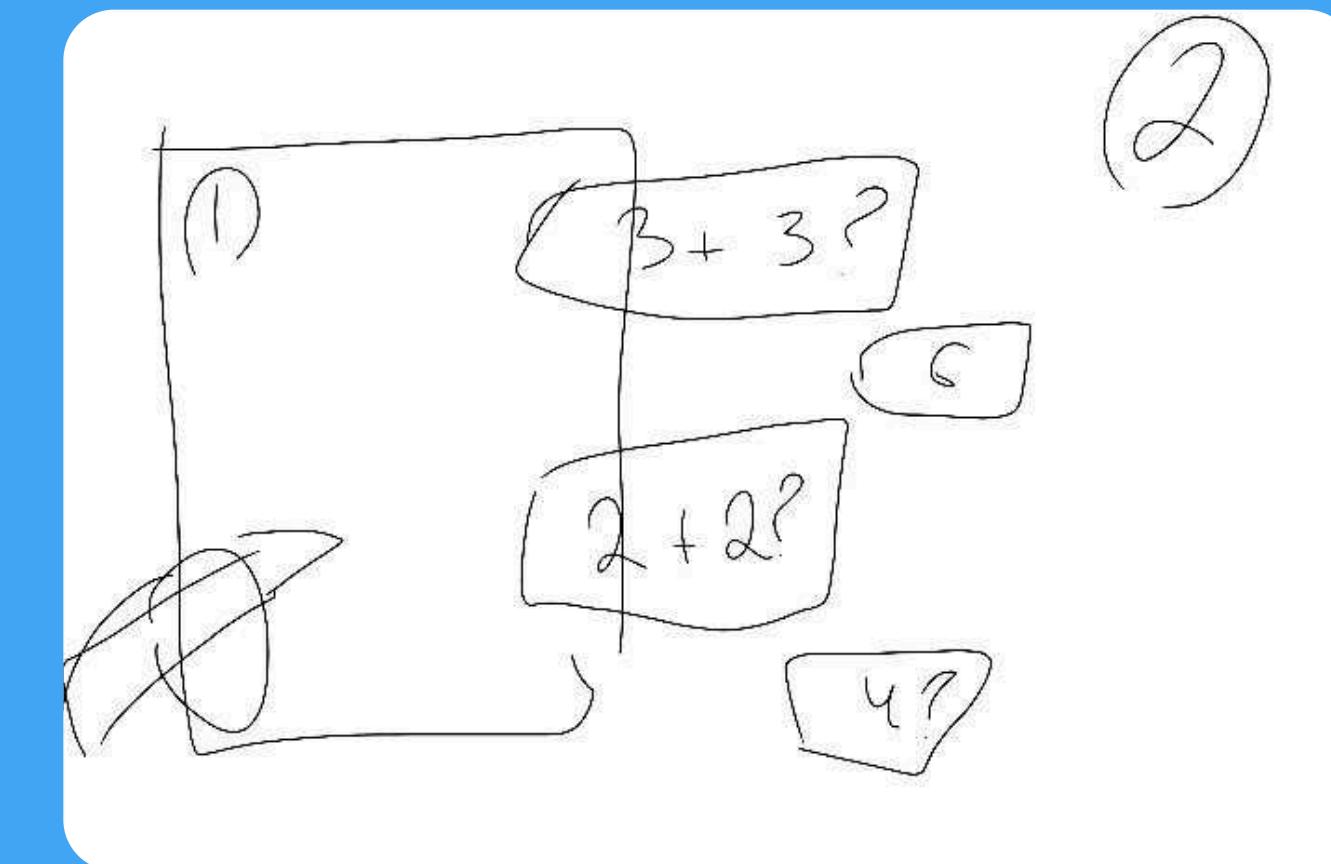
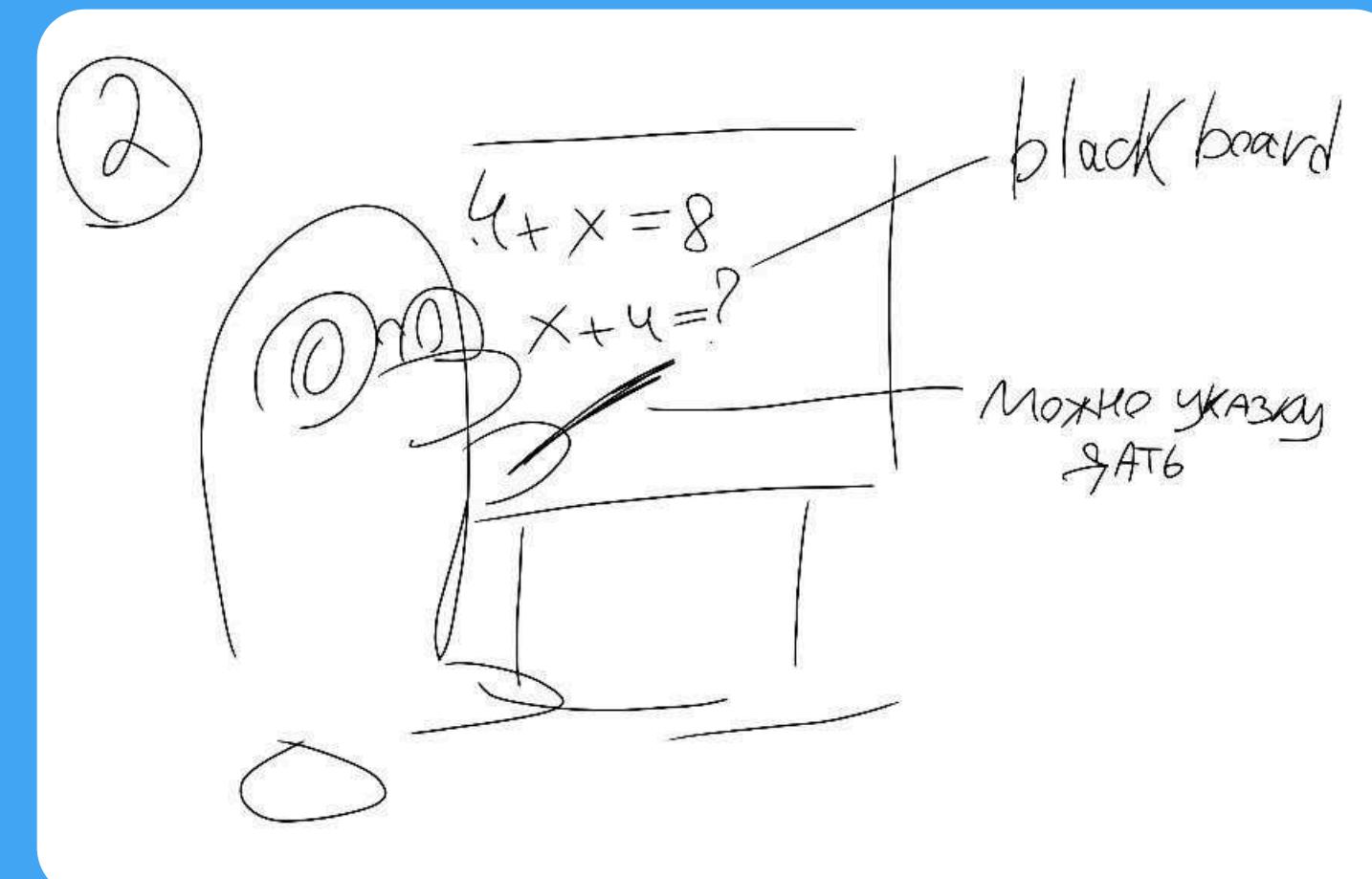
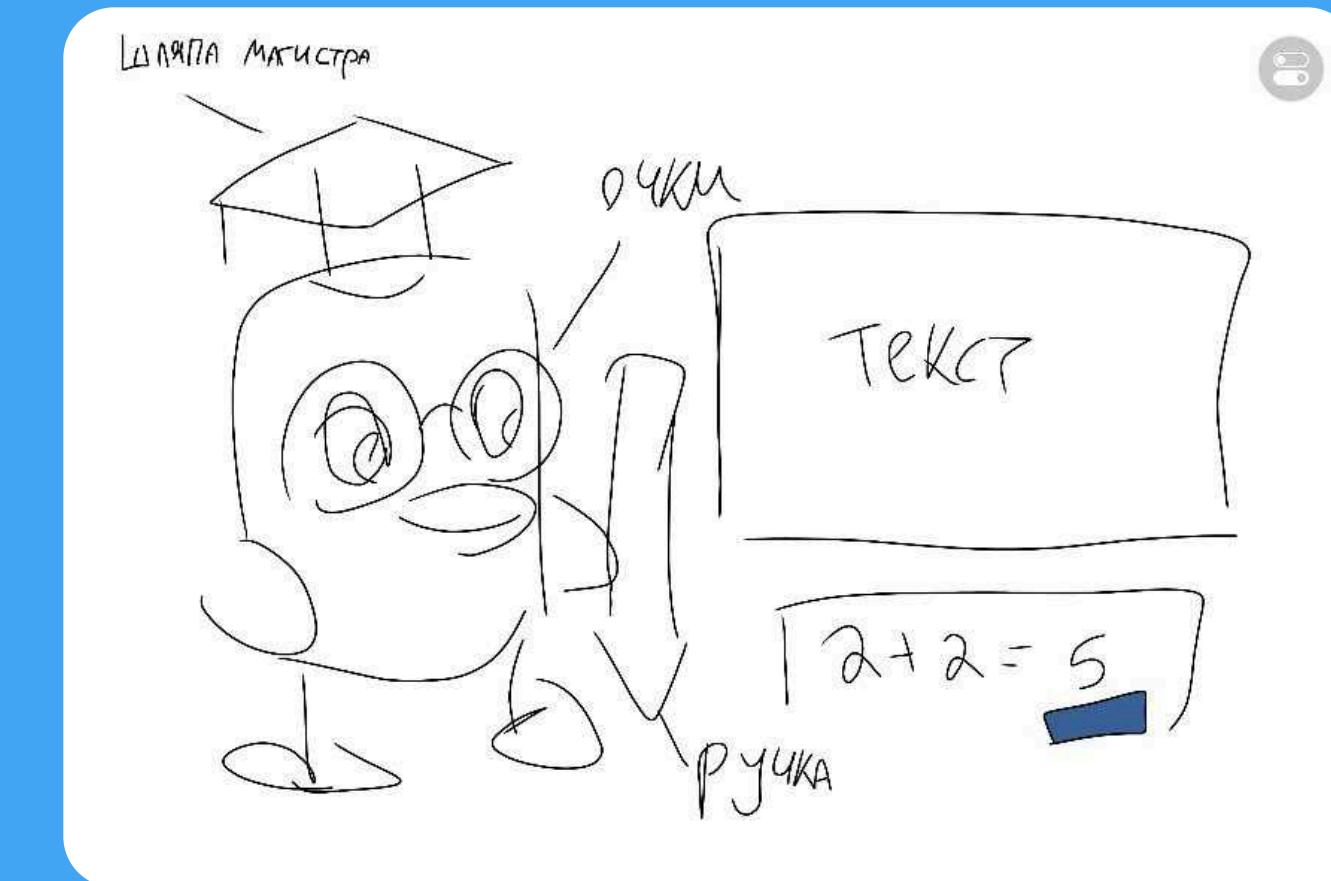
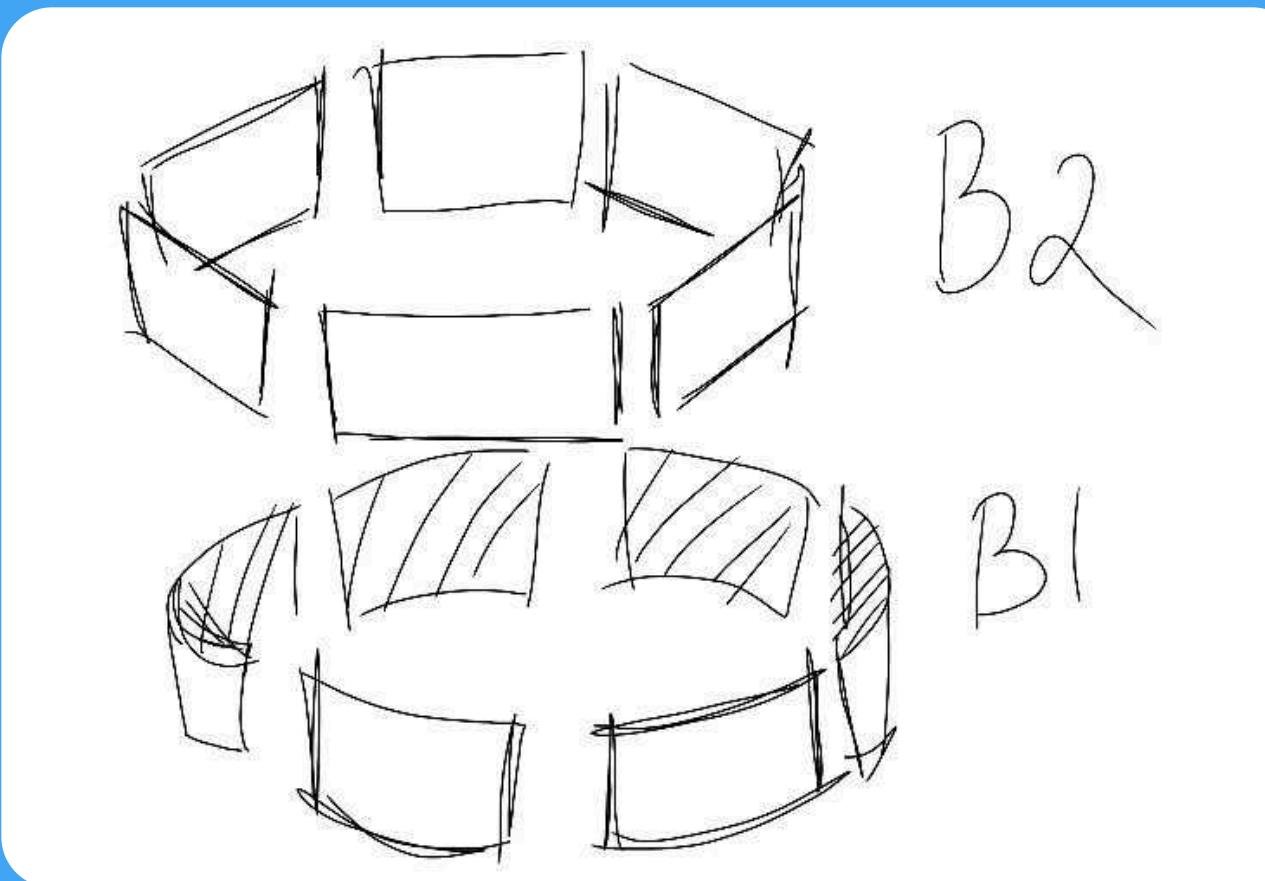


Мати

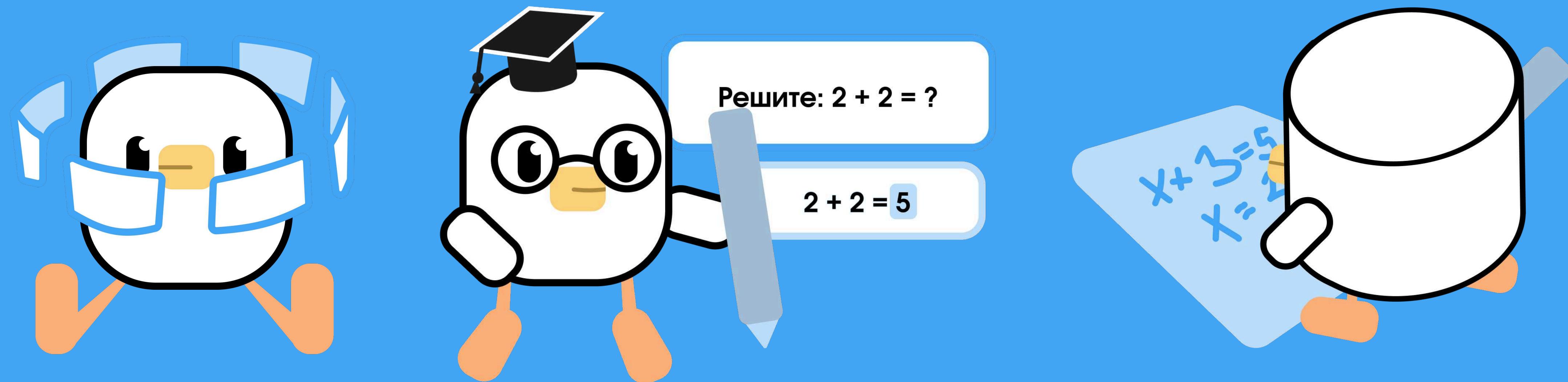


Картина из пинтереста

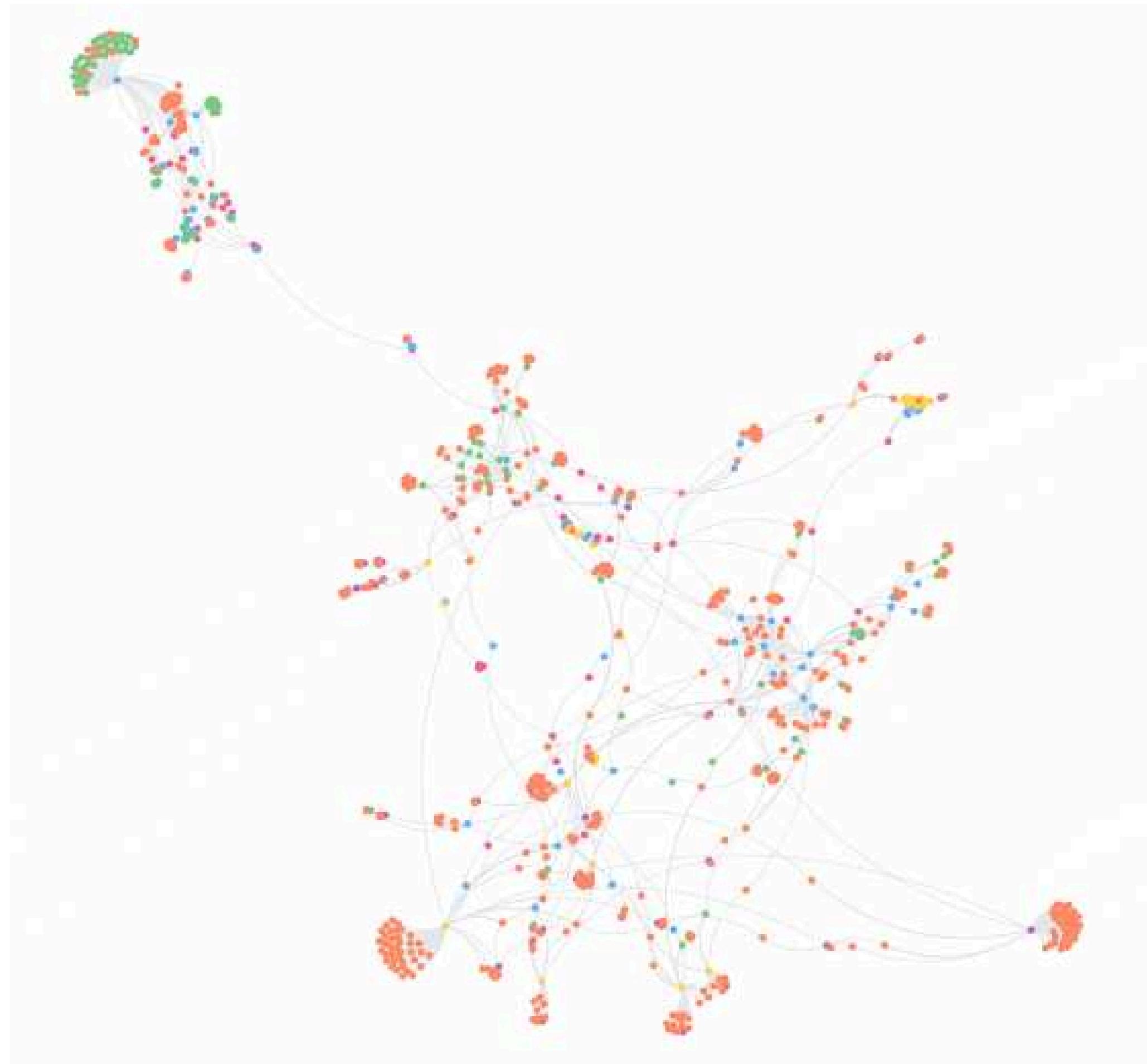
Мати



Мати



Дизайн архитектуры



Бэкенд

Расширяемая DDD архитектура

FastAPI + Pydantic v2 + SQLAlchemy
(async) + SQLite (aiosqlite) + httpx +
OpenRouter (LLM/OCR)

Фронтенд

Лёгкая, модульная архитектура с

чётким разделением

инфраструктуры и фич

react + zustand + axios + framer-motion + react-markdown + tailwindcss + typescript + vite и т.д.

БЭКЕНД

Слои (DDD)

API: роутеры (подключаются в app.main) и Pydantic схемы

Core: ошибки, логирование, security/JWT, DI-хелперы (deps.py)

Domain: сущности и интерфейсы репозиториев

Infrastructure: SQLAlchemy-модели/репозитории/UoW, LLM-роутер + адаптеры, OCR-адаптер, локальное хранилище, CSV-репозиторий задач на pandas, система промптов.

Эндпоинты

Системные:

GET / — проверка доступности сервиса.

GET /health — возвращает статус и версию приложения.

Аутентификация:

POST /api/auth/register — создать нового пользователя.

POST /api/auth/login — войти и получить JWT-токен.

GET /api/auth/profile — получить данные текущего пользователя.

GET /api/auth/stats — вернуть его статистику (решённые задачи, монеты и т. д.).

Темы:

GET /api/themes — список всех тематик задач.

Задачи:

GET /api/tasks — вернуть задачи с фильтрами (по темам, сложности, строке поиска, исключая решённые).

GET /api/tasks/{id} — получить конкретную задачу.

GET /api/tasks/daily — вернуть задачу дня.

GET /api/tasks/{id}/attempts — показать все попытки пользователя по этой задаче.

Решения и помощник:

POST /api/assistant/hint — LLM-подсказка по задаче на основе истории чата.

POST /api/teacher/init — инициализация teacher-режима (LLM-разбор решения).

POST /api/teacher/message — отправить сообщение преподавателю-LLM в режиме разбора.

POST /api/solve — сгенерировать полное решение задачи с пояснениями.

POST /api/check — проверить пользовательское решение, выделить ошибки (через LLMRouter).

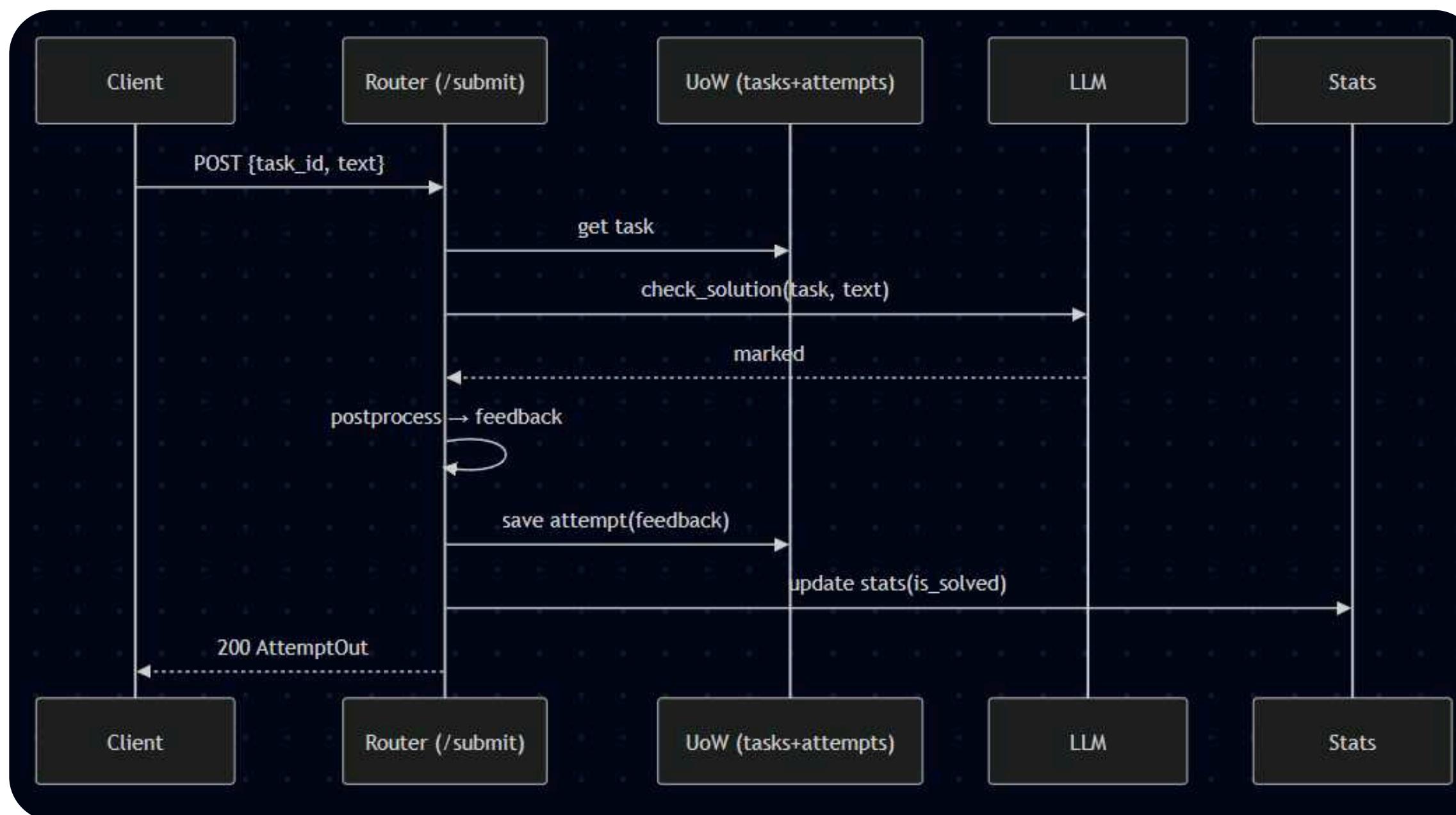
OCR и загрузки:

POST /api/submit/image — принять изображение с решением и распознать текст.

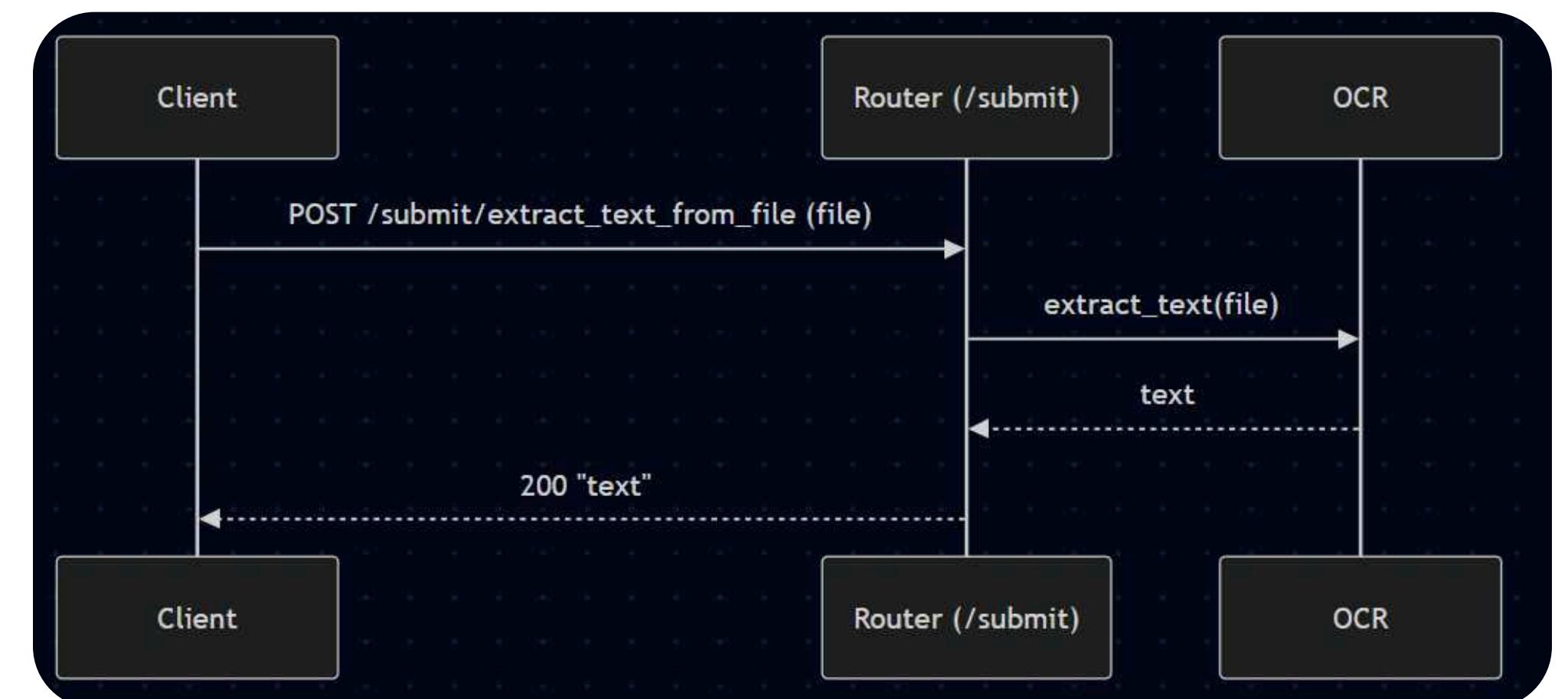
POST /api/submit/pdf — принять PDF и извлечь текст для проверки.

Пример работы эндпоинтов

Отправка решения



OCR



Фронтенд

Структура

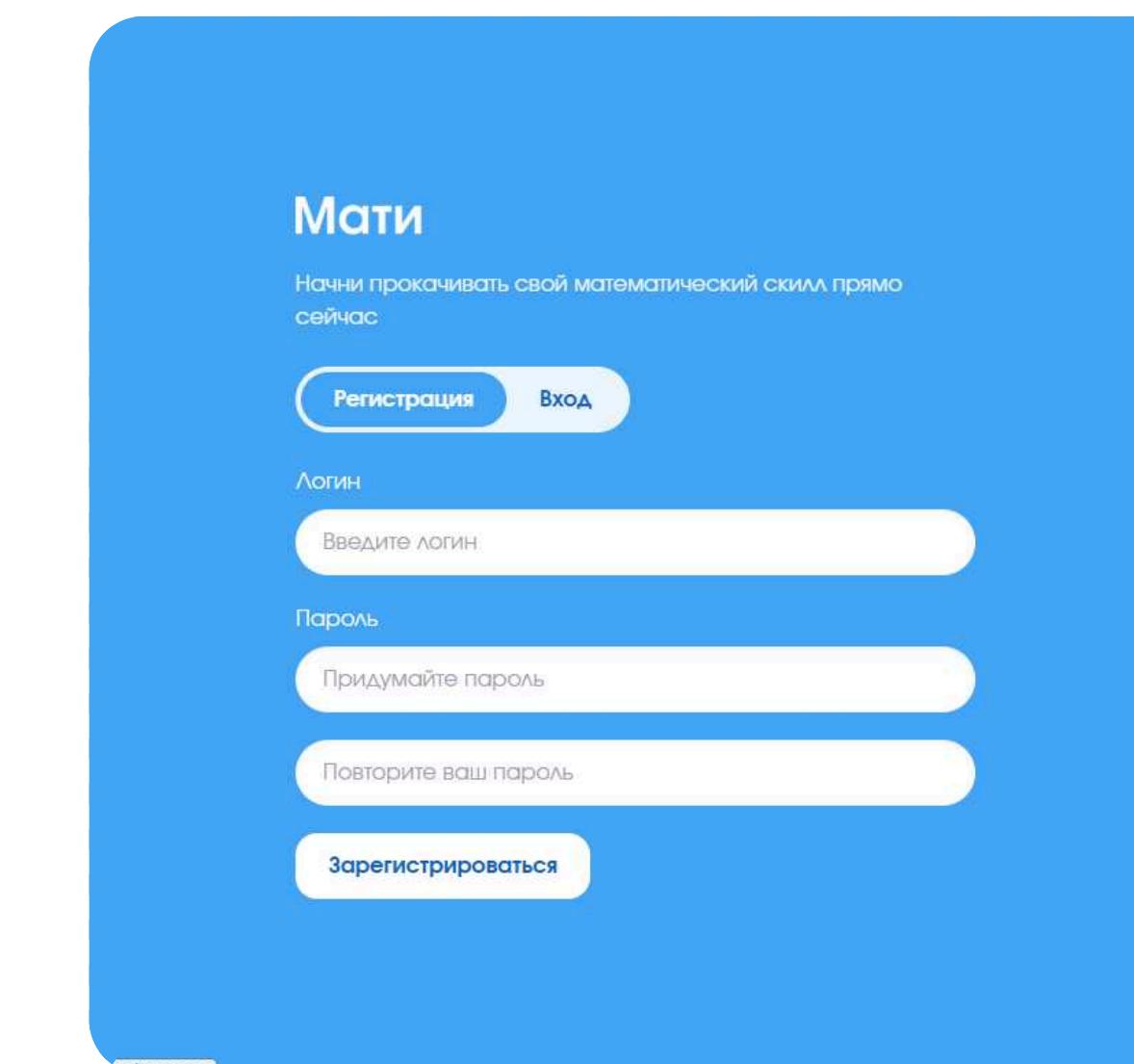
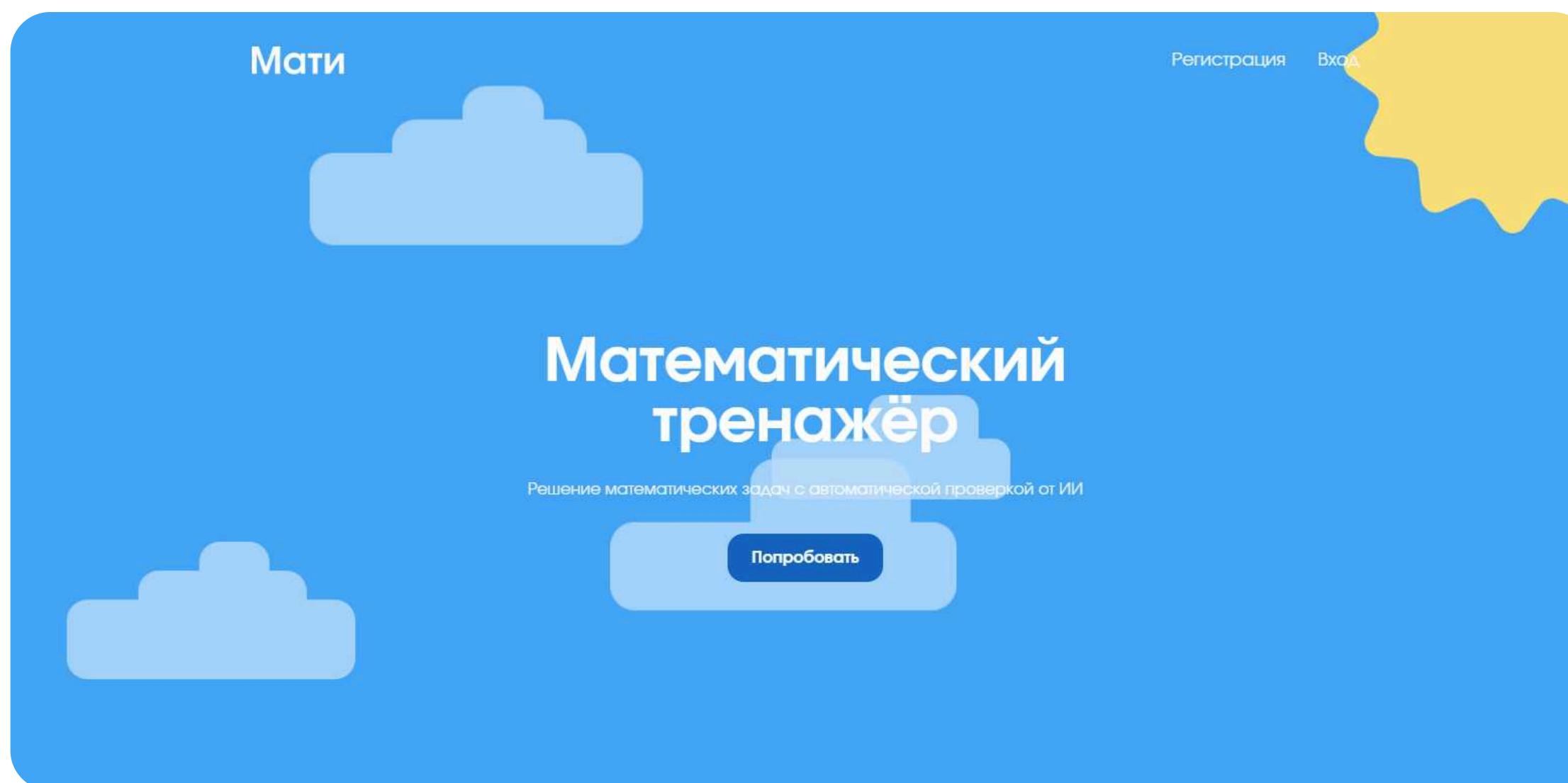
Простой прототип на React + TypeScript + Vite с архитектурой feature-first

Интерфейс лёгкий и отзывчивый: Tailwind, Framer Motion

```
frontend/
  src/
    app/                      # маршрутизация, ProtectedRoute
    components/                # UI-атомы: Button, TextInput, Markdown, CopyButton, MascotEyes, CloudField
    lib/
      api/                    # axios, stream.ts, stream_delta.ts
      store/                  # auth.store.ts (Zustand)
    pages/                    # Auth, Main (Tabs: Tasks/Theory/Daily), Task (Solve/Teach)
    styles/                  # tailwind токены, шрифты
    main.tsx                 # React root
  vite.config.ts, tailwind.config.js, ...
```

Разработка

Главный экран → авторизация



Разработка

Банк задачий → экран выполнения

Мати

Теория Задачи Ежедневная задача 6 droyfl

Темы

- Все задачи
- Мат. анализ
- Теор. вер. и статистика
- теория чисел
- Линейная алгебра
- Дискретная математика
- Оптимизация
- Функ. анализ

Уровни сложности

- Любая сложность
- ★★★ Лёгкий
- ★★★ Средний
- ★★★ Сложный

Поиск задачий

6/1331 Решено

✓ Предел последовательности с косинусом
Вычислить $\lim_{n \rightarrow \infty} \frac{n \cos(n^2 + e^n)}{2^n - 1}$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

★★★ Мат. анализ

Предел разности корней
Вычислить $\lim_{n \rightarrow \infty} (\sqrt{n+1} - \sqrt{n})$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

★★★ Мат. анализ

Задача Мои посылки Решение

Режим решения Режим преподавания

Мат. анализ ★★★

Предел разности корней

Вычислить $\lim_{n \rightarrow \infty} (\sqrt{n+1} - \sqrt{n})$. Ответ запишите в виде десятичной дроби, округлив по правилам математического округления до 3 знаков после запятой.

Решение

Введите решение...

Загрузите файл или перетащите сюда (.png, .jpeg, .pdf, .txt, .md, .tex)

Проверить

AI помощник

Застяг? Спроси у меня подсказку и я тебе помогу!

Shift+Enter — перенос строки

Разработка

Теория и ежедневная задача (приятный бонус)

Мати

Теория Задачи Ежедневная задача 6 droyti

Поиск по темам и урокам Развернуть Свернуть

Теория
Изучай темы блоками и ныряй в уроки в один клик

Математический анализ

- Пределы и непрерывность функций Урок 001 Задач по теме: 0
- Производные и дифференциалы; формула Тейлора Урок 002 Задач по теме: 0
- Интегралы: неопределённый, определённый, несобственные Урок 003 Задач по теме: 0
- Ряды: числовые и степенные; признаки сходимости; ряды Тейлора и Фурье Урок 004 Задач по теме: 0
- Функции многих переменных: градиент, экстремумы, кратные и криволинейные/поверхностные интегралы Урок 005 Задач по теме: 0

Линейная алгебра и геометрия

Мати

Теория Задачи Ежедневная задача 6 droyti

Обновить

Ежедневная задача
Обновляется каждый день в 00:00. До следующей: 07:18:38

10 ноября 2025 г. Мат. анализ ★★★ Лёгкий

Вычисление предела функции

Найти $\lim_{x \rightarrow 2} (8 - 2x^2) / (x^2 + 4x - 12)$. Ответ сократи (используя правила математического округления) до трех знаков после точки

Открыть задачу

Подсказки и метаданные

Когда обновляется?
Новая задача появляется каждый день в 00:00 (по вашему локальному времени).

Сложность
★★★ Лёгкий

Тема
Mat. анализ
Раздел курса: Мат. анализ

До следующей задачи
07:18:38

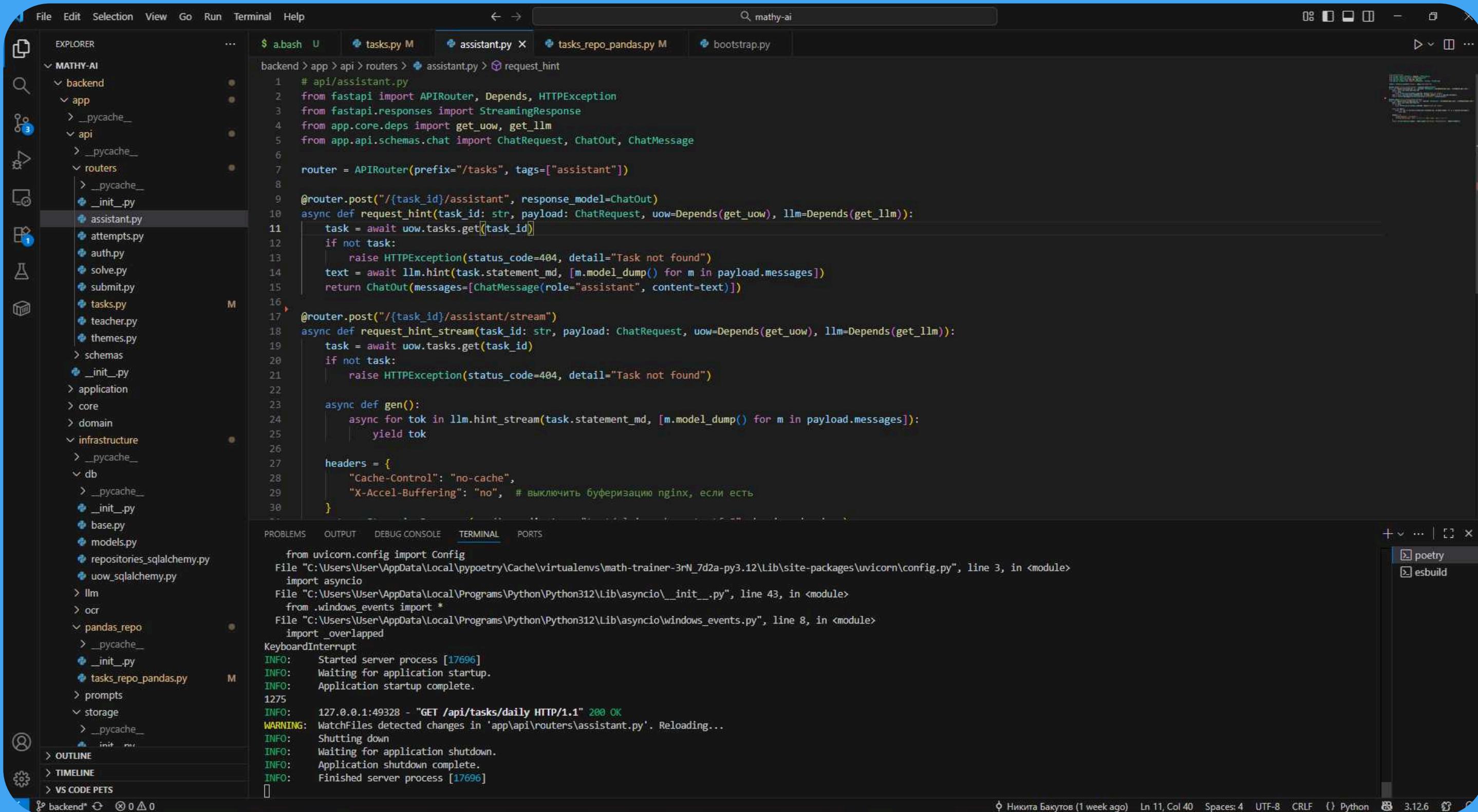
Сложность

Разработка продукта — само по себе очень сложно
Можно запутаться в коде, что-то всегда может сломаться

Инсайты

1

Вайбкодингом пользоваться нужно, но аккуратно.
Всегда нужно самому понимать, что происходит в
проекте



The screenshot shows a dark-themed code editor interface, likely VS Code, displaying Python code for a FastAPI application named 'MATHY-AI'. The code is organized into several files:

- `assistant.py`: Contains the main API router definition.
- `tasks.py`: Contains logic for tasks, including a function `request_hint` which uses `uvicorn` to interact with a database and an LLM.
- `bootstrap.py`: A script used for bootstrapping the application.

The code includes imports from `fastapi`, `uvicorn`, and `app.core.deps`. It defines an API router for tasks, handles task retrieval via `request_hint`, and provides a stream endpoint (`/task_id/assistant/stream`) using `llm_hint_stream`.

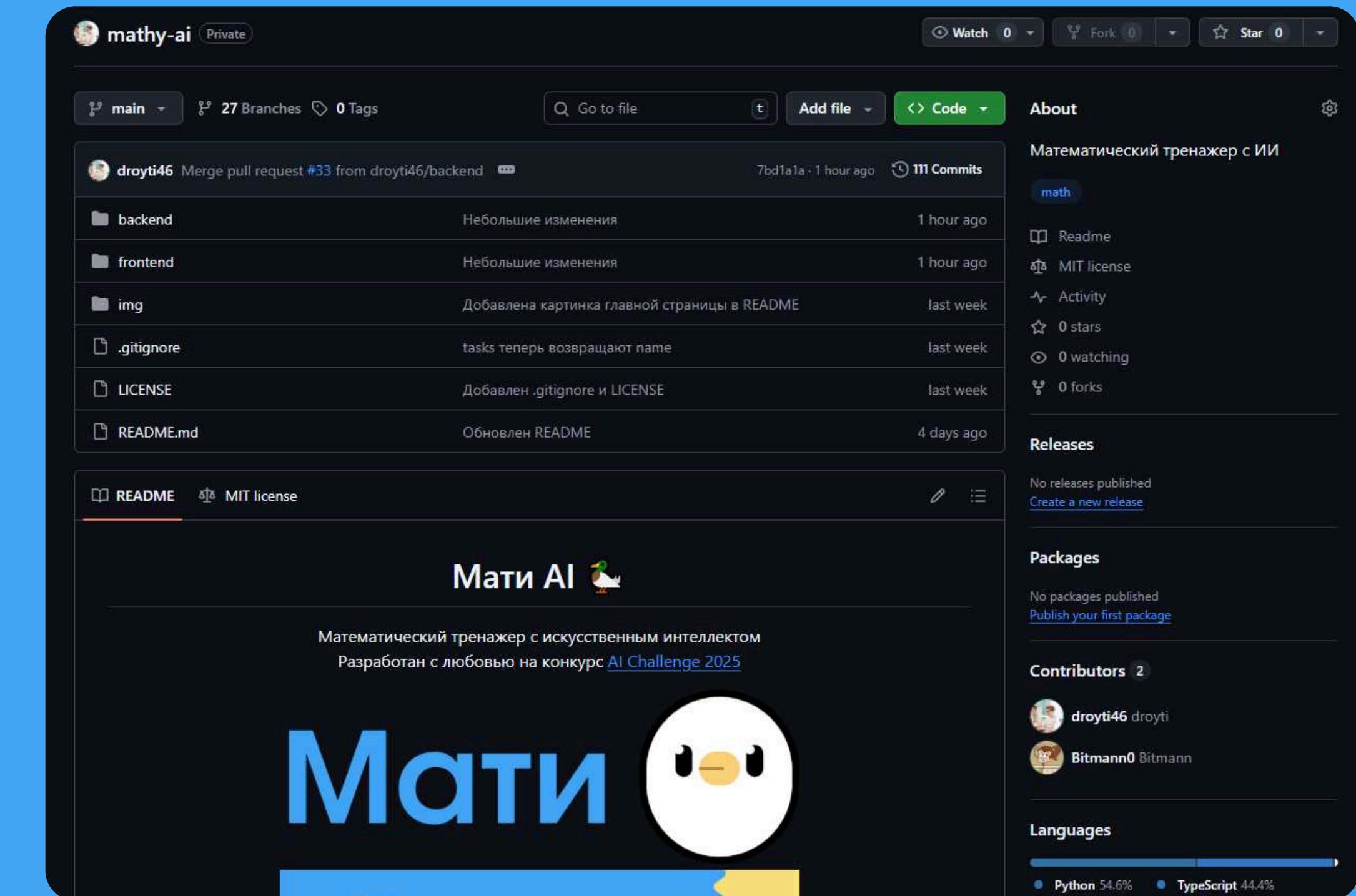
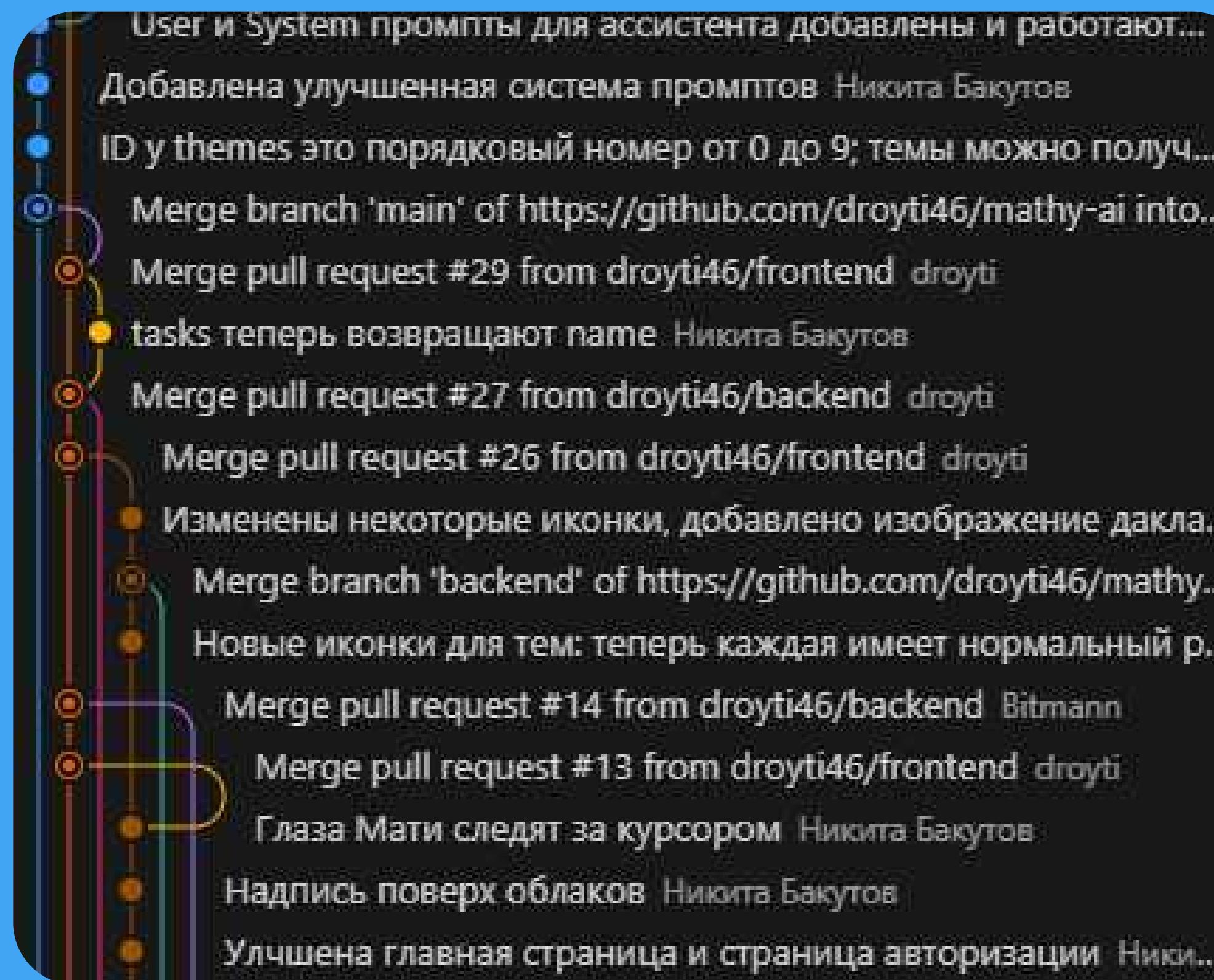
The bottom of the screen shows the terminal output, which logs the server starting at port 149328, waiting for startup, and then shutting down after receiving a shutdown signal.

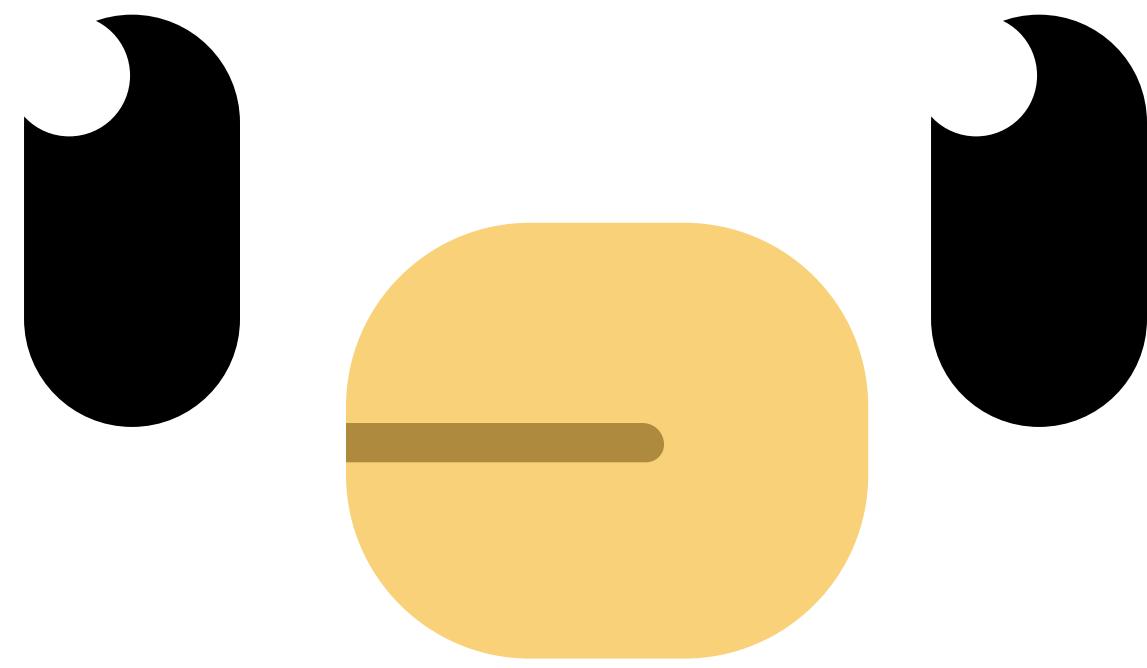
```
INFO: Started server process [17696]
INFO: Waiting for application startup...
INFO: Application startup complete.
1275
INFO: 127.0.0.1:149328 - "GET /api/tasks/daily HTTP/1.1" 200 OK
WARNING: WatchFiles detected changes in 'app/api/routers/assistant.py'. Reloading...
INFO: Shutting down
INFO: Waiting for application shutdown.
INFO: Application shutdown complete.
INFO: Finished server process [17696]
```

Инсайты

2

Система контроля версий — это отлично. Не надо ей пренебрегать





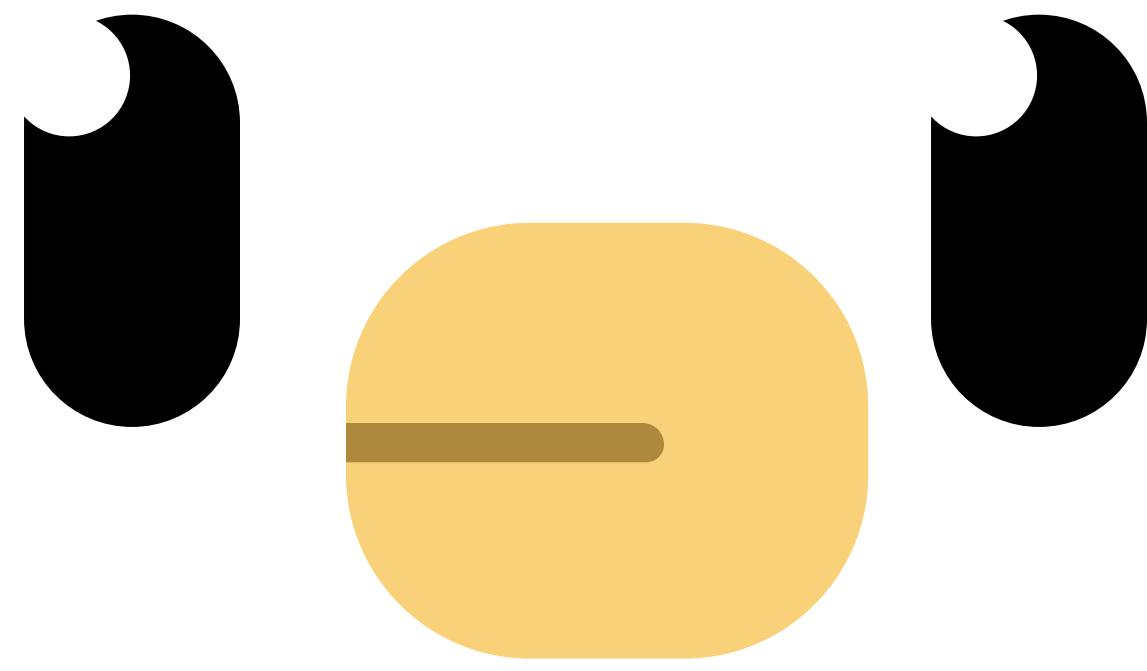
Видение

Дальнейшее виденье проекта



При дальнейшем развитии проекта можно добавить:

- Ежедневные стрики
- Возможность загружать свои задания
- RecSys
- Интеграция математических курсов
- Другие предметы (информатику, химию, физику и т.д.)



Спасибо за внимание

Команда



Андрей Четверяков
ML Engineer,
Backend



Никита Бакутов
ML Engineer, Full-Stack,
Designer



Мария Лабецкая
Frontend



Алина Кудряшова
Frontend

Мати



Сидим
не рыпаемся



AI Challenge



ЦЕНТРАЛЬНЫЙ
УНИВЕРСИТЕТ