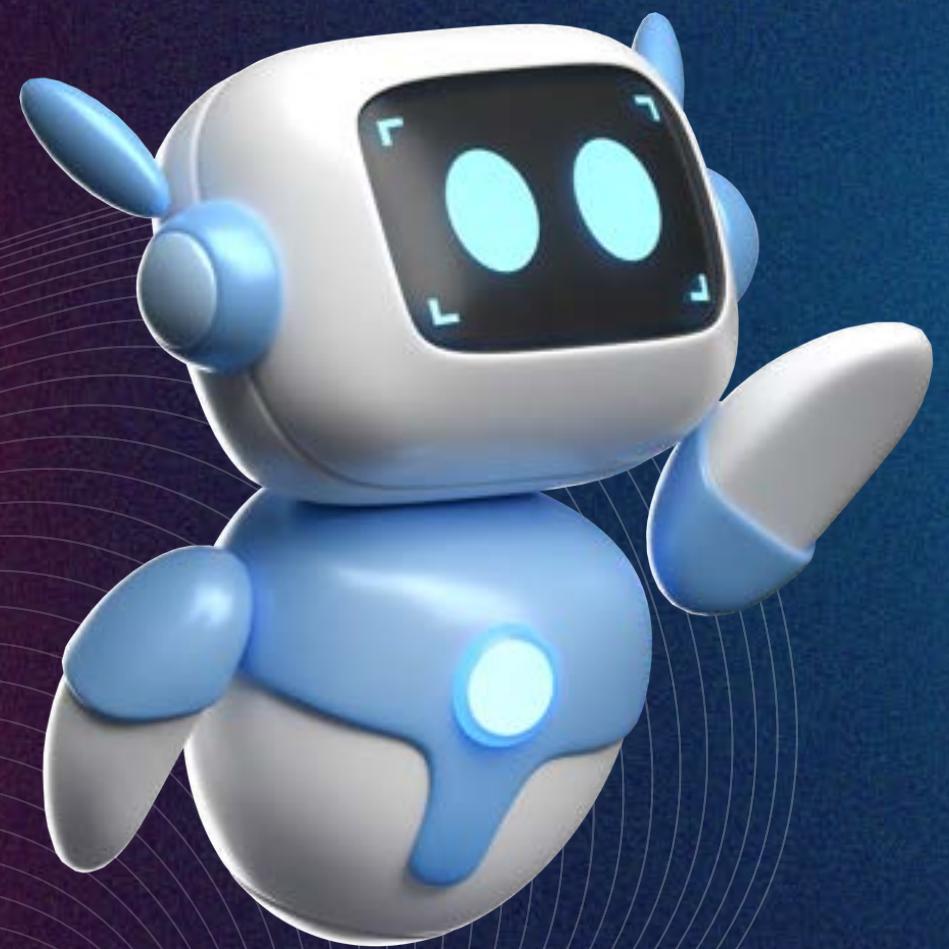


SOLVING ONET MATH QUESTIONS WITH SMALL LANGUAGE MODELS



Presented by

ชุมอุ่นเปเปอร์ NLP

6430001121

Kanawat Vilasri

6430038421

Jakrin Sukonpongao

6432114821

Phichet Phuangrot

6432147521

Rossukhon Rukkiattikhun

6532163621

Veerawat Vanamonthon

AGENDA

- Topic
- Dataset
- Model
- Resource
- Paper
- Methodology
- Result
- Q&A

TOPIC

We aim to build small language models that can solve ONET math exam questions at a level similar to the average Grade 12 student in Thailand.

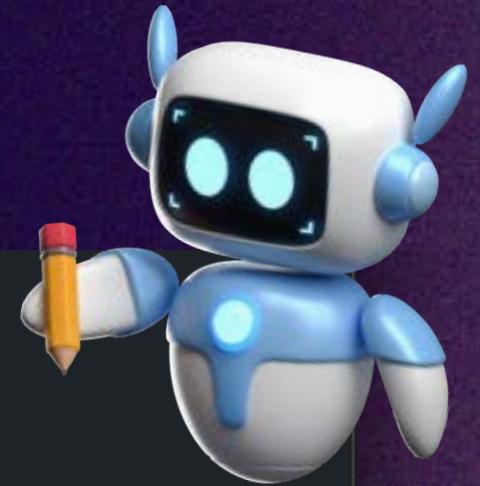
We use a dataset of over 180 past ONET math questions for **training**. For **testing**, we use a separate set of 16 old ONET math questions that the models haven't seen before.

We focus on **3 small language models**:

1. facebook/xglm-564M
2. facebook/xglm-2.9B
3. scb10x/llama3.2-typhoon2-3b-instruct

TOPIC

We evaluate model performance using **accuracy** and compare the scores with national rankings from [\[thaiexamjs/onet.html\]](#)



LLM Performance on Thai O-NET Tests

This dashboard showcases how different AI models perform on O-NET standardized tests.

[← Back to Index](#)

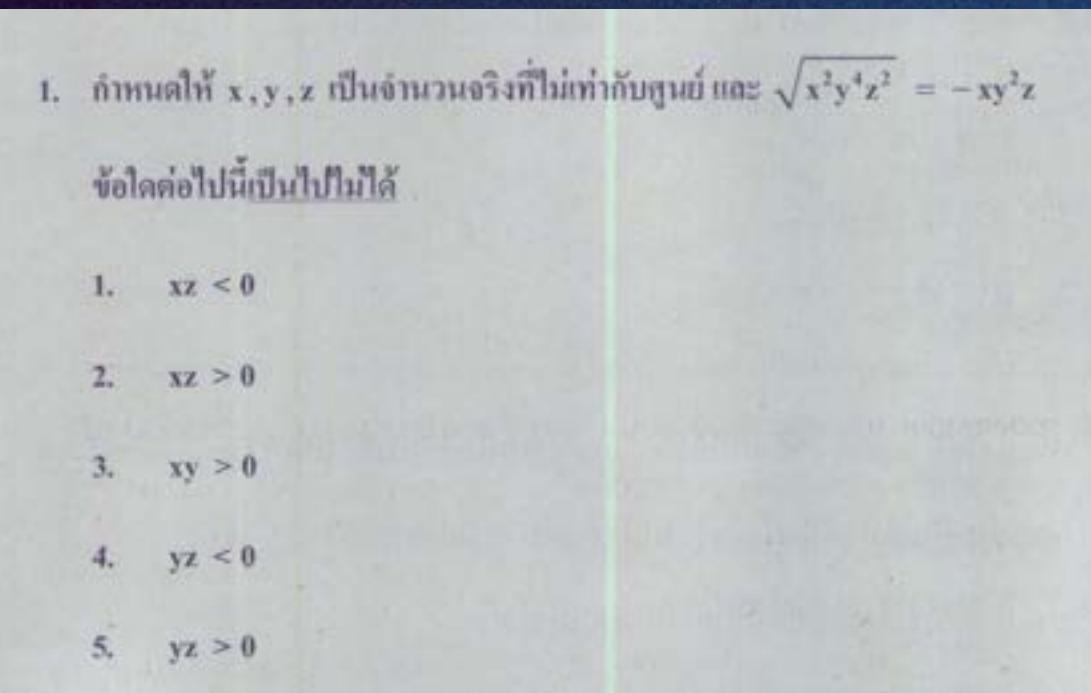
[View on GitHub](#)

Overall ranking

Model	Cost	Thai	Social	Science	Math	Overall	Acc
claude-3-7-sonnet-20250219[thinking=16k]	\$159.13	51/63	56/63	18/20	15/16	140/162	86.42%
M6 students average (adjusted from stats published by NIETS)		29/63	23/63	6/20	3/16	62/162	38.02%

DATASET

ONET math M.6 since 2558-2567



ocr by

Gemini 2.0
Flash

prompt to return in Latex form

test	years	problem	question	choice 1	choice 2	choice 3	choice 4	choice 5	choice_ans	num_ans
onet	2558	2	ถ้า \$a < b\$ และ $\frac{1}{a} < \frac{1}{b}$ แล้ว \$a < b\$ และ \$c < d\$ แล้ว $\sqrt{(a+b)^2} = a+b $ \$ a+b = a + b							4
onet	2558	3	จำนวนจริง $\sqrt{84 + 18\sqrt{3}}$ มีค่า $\$4 + 3\sqrt{3}$ $\$5 + 2\sqrt{3}$ $\$6 + 2\sqrt{3}$ $\$9 + \sqrt{3}$ $\$10 + \sqrt{3}$							4
onet	2558	4	ถ้า \$a = -5\$ และ \$b = 8\$ และ $\sqrt{6}(a^2b)$	10	-10	20	-15	-40	1	
onet	2558	5	ถ้า $\sum_{i=1}^5 x_i = -10$ และ $\sum_{i=1}^5$	11	12	13	14	15	2	
onet	2558	6	ถ้า $x = 1 + \sqrt{3}$ และ $\frac{x^4}{(x^2-1)^2} = 1 + \sqrt{3}$							5
onet	2558	7	ถ้า \$A = \{x \mid \mathbb{R} \mid x+1 < 0\} \cap B = \{0\}\$							3
onet	2558	8	ถ้าหน้า "เหลี่ยม" เป็นเหลี่ยม หน้าที่ออกก้าวกระโดด หน้าที่ไม่ออกก้าวกระโดด หน้าที่ออกก้าวกระโดด หน้าที่ไม่ออกก้าวกระโดด หน้าที่ไม่ออกก้าวกระโดด หน้าที่ไม่ออกก้าวกระโดด							1
onet	2558	9	ถ้า $f(x) = x + x $ และ x ไม่ต้องไถูก							4
onet	2558	11	ถ้า $x + y = 1$ และ x ค่าต่ำสุดของ $x^2 + 2xy$							1
onet	2558	12	ถ้า $ x+1 = 3$ และ x มีค่าต่ำสุดเท่ากับ -5							1
onet	2558	13	ถ้าหน้า "ABC" เป็นรูปสามเหลี่ยมแนบในรูป $\frac{55}{\sqrt{3}}$ หน่วย	10 หน่วย	15 หน่วย	$10\sqrt{3}$ หน่วย	20 หน่วย			2
onet	2558	16	โดยที่อนันต์เป็นไปในแนวที่ต้องตัดคราว 96 ครั้ง $\frac{1}{100} < t < 25$							2

then checking by human



reasoning by
Deepseek-R1-Distill-Qwen1.5B

DATASET

ONET math M.6 since 2558-2567



DATASET

Datasets: [TaengooTV/math_onet](#) like 0

Tasks: Question Answering Modalities: Tabular Text Formats: csv Languages: Thai Size: <1K Tags: math Libraries:

[Dataset card](#) [Data Studio](#) [Files and versions](#) [Community 1](#) [Settings](#)

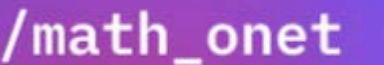
Dataset Viewer

Split (1)
train · 191 rows

Search this dataset

test	years	problem	question	choice 1	choice 2	choice 3
string · classes	int64	int64	string · lengths	string · lengths	string · lengths	string · lengths
1 value	2.56k	2.57k	1	40	24	553
onet	2,558	2	2	93	1	81
onet	2,558	3	3	93	1	81
onet	2,558	4	4	93	1	81
onet	2,558	5	5	93	1	81
onet	2,558	6	6	93	1	81
onet	2,558	7	7	93	1	81

< Previous 1 2 Next >

 [TaengooTV](#)
 [/math_onet](#)

 [huggingface.co](#)

TaengooTV/math_onet · Datasets at Hugging Face
We're on a journey to advance and democratize artificial intelligence through open source and open science.

 [huggingface](#)

TEST SET

ONET math M.6 2564 - 16 problems

Datasets: [openthaigpt/openthaigpt_eval](#) [Follow](#) [OpenThaiGPT](#) 95

Tasks: [Question Answering](#) Languages: [Thai](#) Size: [n<1K](#) Tags: [exam](#) License: [apache-2.0](#)

dtinth / thaixamjs

Code Issues Pull requests Actions Projects Security Insights

thaixamjs Public

Sponsor Watch Fork Star

main 3 Branches 0 Tags Go to file Add file Code About

dtinth add gemini-2.5-pro-preview-05-06 4e04411 · 3 days ago 70 Commits

.github Project revamp, 2025-04 (#5) 3 weeks ago

openthaigpt_eval Project revamp, 2025-04 (#5) 3 weeks ago

scripts Project revamp, 2025-04 (#5) 3 weeks ago

src add gemini-2.5-pro-preview-05-06 3 days ago

thai_exam first commit 3 months ago

.gitignore Add M6 average score to exam reports and set up GitHub... 3 months ago

.npmrc first commit 3 months ago

CLAUDE.md Project revamp, 2025-04 (#5) 3 weeks ago

README.md Project revamp, 2025-04 (#5) 3 weeks ago

bun.lock add gemini-2.5 (no thinking) variant 3 weeks ago

mise.toml use bun v1.2.4 2 months ago

package.json add gemini-2.5 (no thinking) variant 3 weeks ago

snapshot.json.br add gemini-2.5-pro-preview-05-06 3 days ago

Compare how different AI models perform on Thailand's standardized exams

dtinth.github.io/thaixamjs/

Readme Activity 27 stars 1 watching 0 forks

Report repository

No releases published

Sponsor this project

dtinth Thai Pangakulyanont

ko-fi.com/dtinth

Learn more about GitHub Sponsors

MODEL

We chose **3 small language models** that can understand Thai:

1. **facebook/xglm-564M** [xglm-564M]

- A multilingual language model with 564 million parameters.
- Introduced in the paper Few-shot Learning with Multilingual Language Models.

2. **facebook/xglm-2.9B** [xglm-2.9B]

- A multilingual language model with 2.9 billion parameters.
- Introduced in the paper Few-shot Learning with Multilingual Language Models.

3. **scb10x/llama3.2-typhoon2-3b-instruct** [typhoon2-3b-instruct]

- A Thai instruction-tuned model with 3 billion parameters.
- Based on Llama3.2-3B.
- Designed to follow instructions in Thai more effectively.

RESOURCE

We used LANTA (Large-scale AI Training Architecture).

LANTA provided:

- Fast and stable training, even with large datasets
- High-performance **multi-GPU** infrastructure
- Better speed and reliability compared to traditional cloud services

This allowed us to train and evaluate our models without system crashes or delays (**generation time: days → hours**).



S1: SIMPLE TEST-TIME SCALING

Budget forcing technique

How many r in raspberry?

Question

Let's break down the process of counting the letter 'r' in the word "raspberry" ...

- * First letter: 'r' - This is an 'r', count = 1.
- * Second letter: 'a' - Not an 'r', count remains 1 ...
- * Sixth letter: 'e' - Not an 'r', count remains 1.
- * Seventh letter: 'r' - This is an 'r', count = 2.
- * Eighth letter: 'y' - Not an 'r', count remains 2 ...

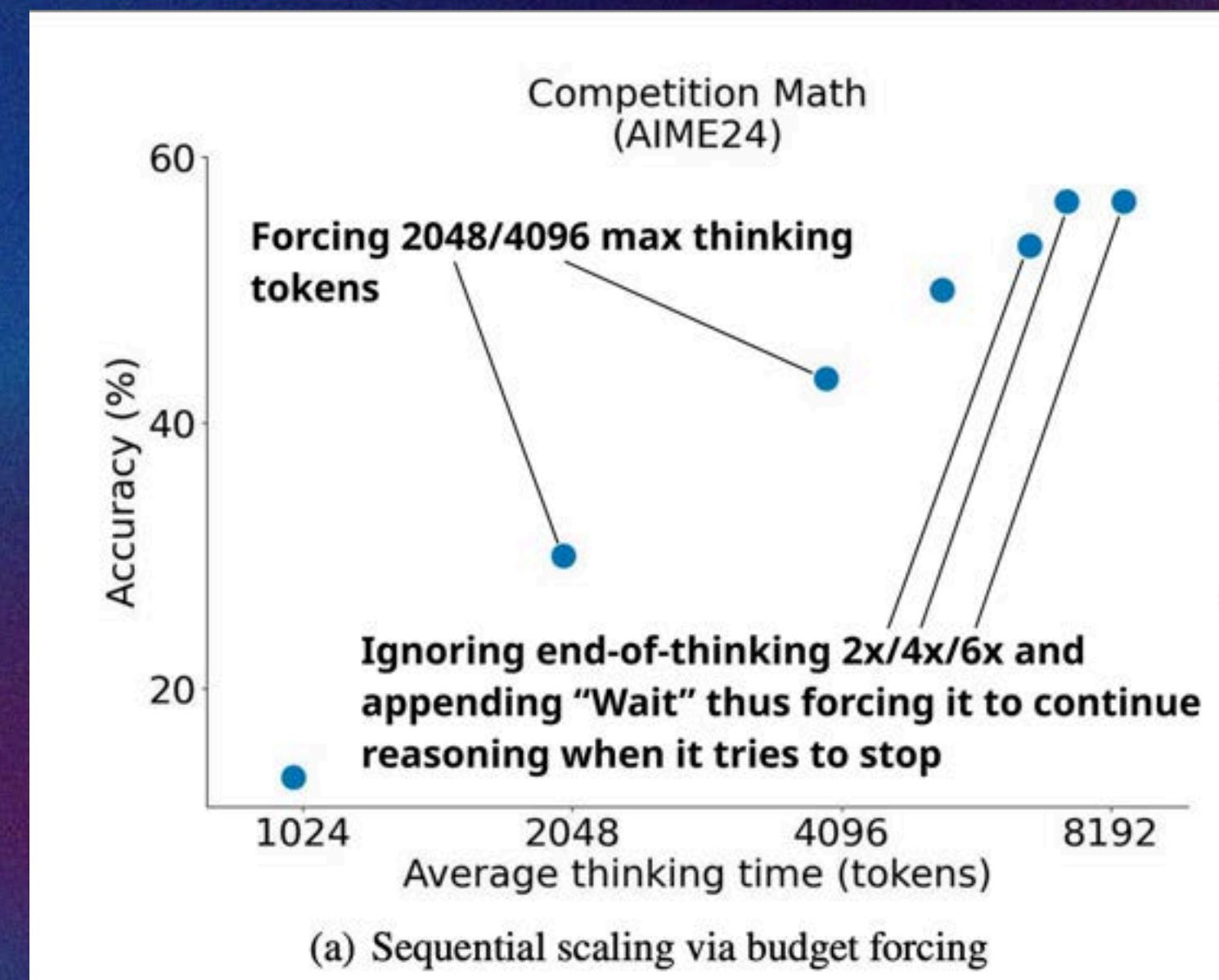
The number of 'r's in "raspberry" is 2.

Wait, let's re-read the question carefully. It asks "How many r in raspberry?" ... * r - a - s - p - b - e - r - r - y ... * First 'r' ... * Second 'r' ... * Third 'r' ... Count = 3 ...

Reasoning trace

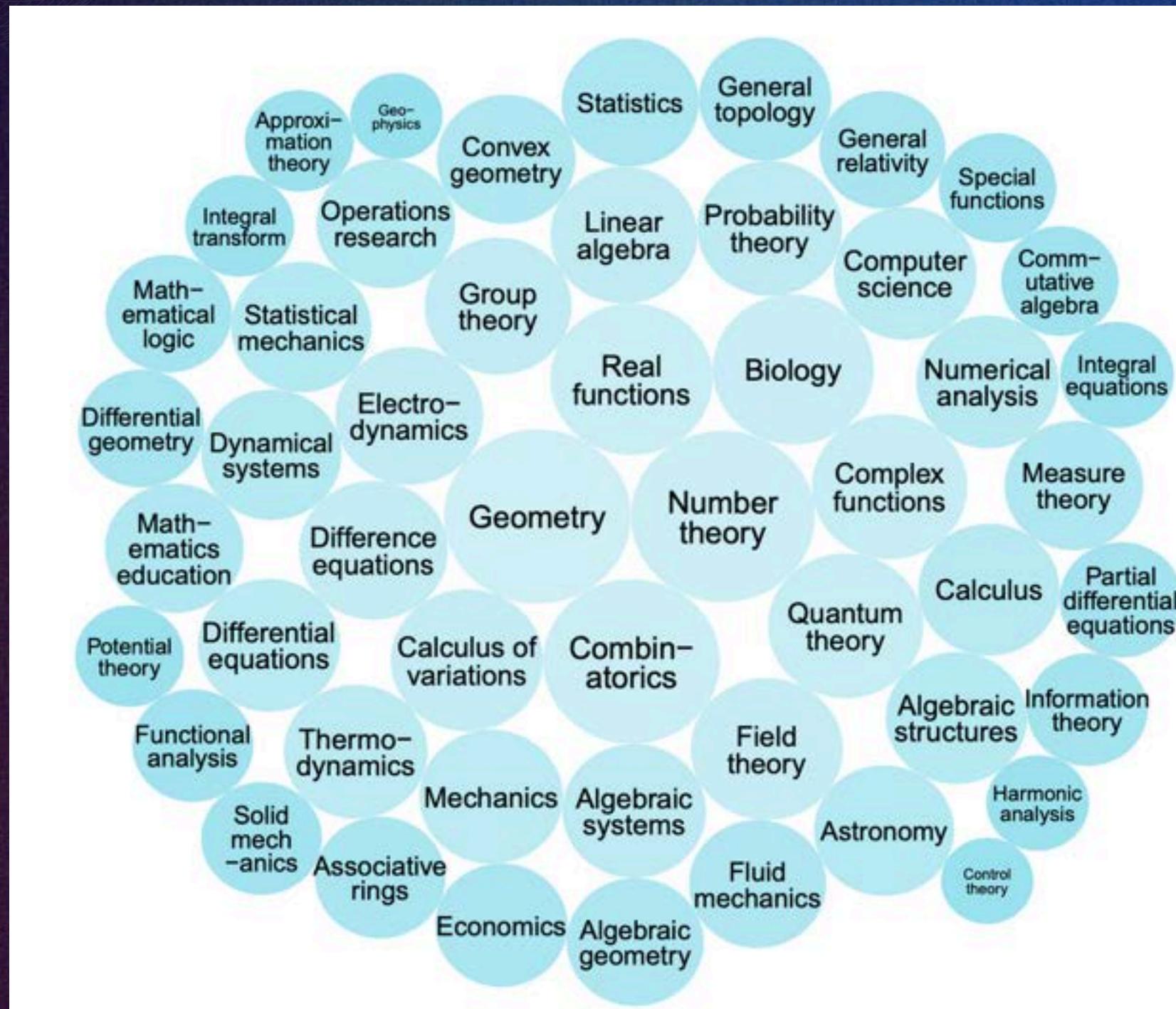
My initial answer of 2 was incorrect due to a quick reading of the word. **Final Answer:** The final answer is **3**

Response



S1: SIMPLE TEST-TIME SCALING

S1K data



Quality

- filter out low-quality examples

Difficulty

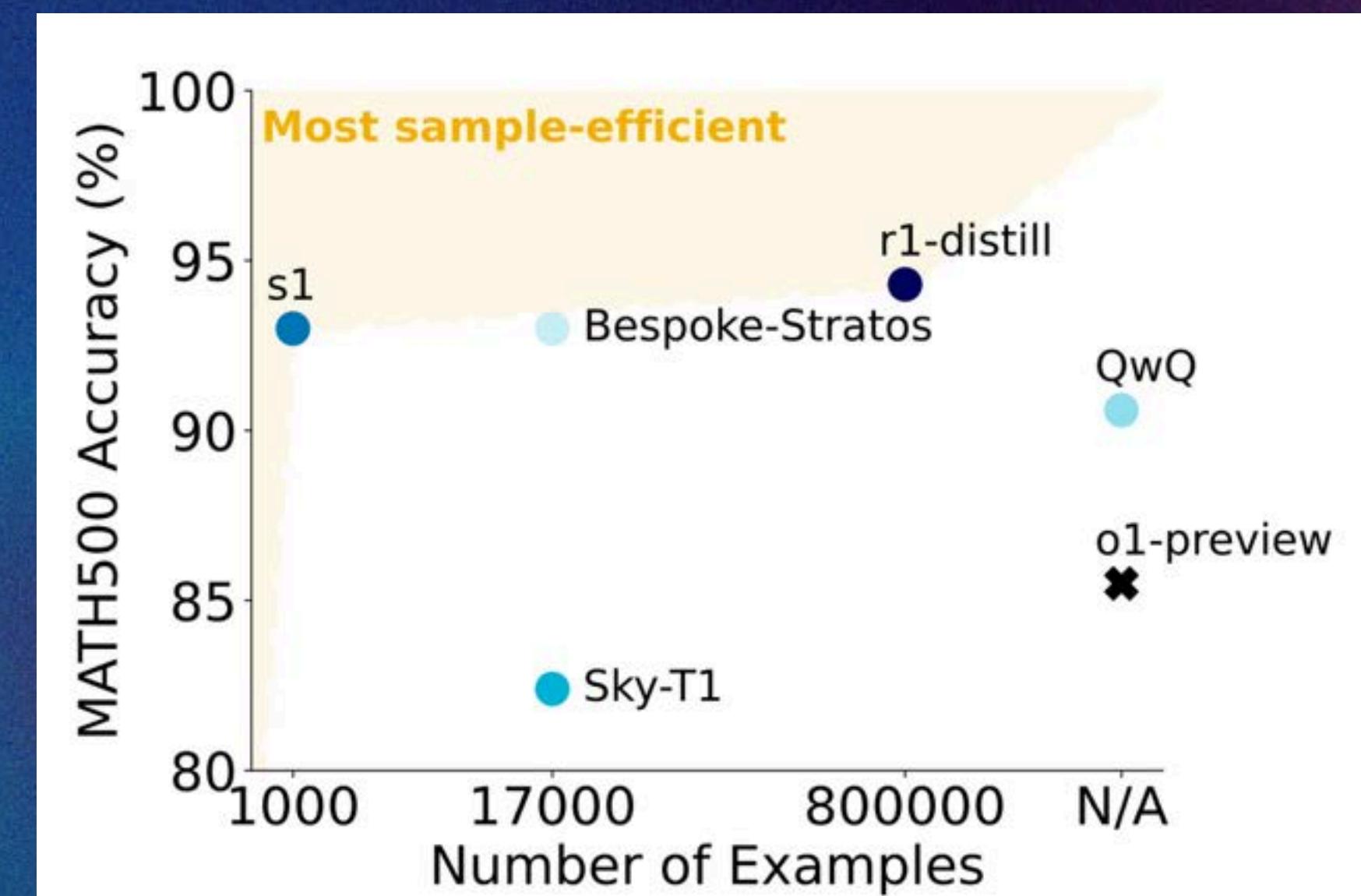
- use two indicators: model performance and reasoning trace length

Diversity (50 domains)

- classify questions into domains using Claude 3.5 Sonnet based on the Mathematics Subject Classification (MSC) system
 - also includes other sciences such as biology, physics, and economy

S1: SIMPLE TEST-TIME SCALING

Model	# ex.	AIME 2024	MATH 500	GPQA Diamond
API only				
o1-preview	N.A.	44.6	85.5	73.3
o1-mini	N.A.	70.0	90.0	60.0
o1	N.A.	74.4	94.8	77.3
Gemini 2.0	N.A.	60.0	N.A.	N.A.
Flash Think.	N.A.			
Open Weights				
Qwen2.5-32B-Instruct	N.A.	26.7	84.0	49.0
QwQ-32B	N.A.	50.0	90.6	54.5
r1	»800K	79.8	97.3	71.5
r1-distill	800K	72.6	94.3	62.1
Open Weights and Open Data				
Sky-T1	17K	43.3	82.4	56.8
Bespoke-32B	17K	63.3	93.0	58.1
s1 w/o BF	1K	50.0	92.6	56.6
s1-32B	1K	56.7	93.0	59.6



S1: SIMPLE TEST-TIME SCALING

Table 4. Budget forcing extrapolation ablations. We compare ignoring the end-of-thinking delimiter twice and appending none or various strings.

Model	AIME 2024	MATH 500	GPQA Diamond
No extrapolation	50.0	93.0	57.6
2x without string	50.0	90.2	55.1
2x “Alternatively”	50.0	92.2	59.6
2x “Hmm”	50.0	93.0	59.6
2x “Wait”	53.3	93.0	59.6

Budget forcing extrapolation ablations

METHODOLOGY

We tested 3 different training methods to improve how well the models solve ONET math questions. The table below shows which model used which method:

Training Method \ Model Used	facebook/xglm-564M	facebook/xglm-2.9B	scb10x/llama3.2-typhoon2-3b-instruct
1. Reasoning model training	Yes	Yes	
2. Basic fine-tuning	Yes		
3. Budget forcing method			Yes

METHODOLOGY : REASONING MODEL

Create reasoning dataset distilled from

- deepseek-ai/DeepSeek-R1-Distill-Qwen-1.5B
 - utilizing vLLM library for fast inference
- prompt:
 - reference from model card

Solve the following problem.

[problem + choices]

Please reason step by step, and put your final

answer within boxed{}.

`<think>`

- max_tokens: 32768

METHODOLOGY : REASONING MODEL

Data Preparation: Format each question so the model generates step-by-step reasoning followed by the final answer.

[question]

ตัวเลือก:

[choices]

[task]

ans:



[question]

ตัวเลือก:

A. ...

B. ...

C. ...

D. ...

E. ...

กรุณารอทิ้งเวลา 10 วินาที ให้ระบบคำนวณและระบุคำตอบที่ถูกต้องที่ที่ทิ้งคำอธิบาย (ในรูปแบบ: คำตอบคือ: **b**)

ans:

METHODOLOGY : REASONING MODEL

Evaluating Pre-trained Model Reasoning: Test how well the pre-trained model can generate logical reasoning based on the input prompt.

กำหนดให้ a , b , c และ d เป็นจำนวนจริงใดๆ ข้อใดต่อไปนี้เป็นจริง ตัวเลือก: A. ถ้า $a < b$ และ $\frac{1}{a} > \frac{1}{b}$
B. ถ้า $a < b$ และ $a^2 < b^2$ C. ถ้า $a < b$ และ $c < d$ และ $ac < bd$ D. $\sqrt{(a+b)^2} = |a+b|$ E. $|a+b| = |a| + |b|$ กรุณาอธิบายเหตุผลอย่างเป็นขั้นตอนและระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: ...) ans: ค่าตอบคือ: ...
correct ans: 4.0

จำนวนจริง $\sqrt{84 + 18\sqrt{3}}$ มีค่าเท่าใด ตัวเลือก: A. $4 + 3\sqrt{3}$ B. $5 + 2\sqrt{3}$ C. $6 + 2\sqrt{3}$ D. $9 + \sqrt{3}$ E. $10 + \sqrt{3}$ กรุณาอธิบายเหตุผลอย่างเป็นขั้นตอนและระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: ...) ans: A. $4 + 3\sqrt{3}$
correct ans: 4.0

ถ้า $a = -5$ และ $b = 8$ และ $\sqrt{6(a^2b)}\sqrt{6(a^4b)}$ มีค่าเท่าใด ตัวเลือก: A. 10 B. -10 C. 20 D. -15 E. -40 กรุณาอธิบายเหตุผลอย่างเป็นขั้นตอนและระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: ...) ans: A. 10 B. -10 C. 20 D. -15
correct ans: 1.0

Observation: The pre-trained model often defaults to selecting choice (A) or copies the example answer from the prompt, indicating limited reasoning without fine-tuning.

METHODOLOGY : REASONING MODEL

Fine-Tuning the Model: The goal is to train the model to generate step-by-step reasoning and the correct answer using formatted data that includes reasoning examples from DeepSeek.

```
tokenizer = XGLMTokenizer.from_pretrained("facebook/xglm-2.9B")
model = XGLMForCausalLM.from_pretrained("facebook/xglm-2.9B")
```

PEFT - LoRA

- rank - 4
- lora_alpha=32
- lora_dropout=0.05
- target_modules=["q_proj", "v_proj","k_proj","out_proj"]

METHODOLOGY: REASONING MODEL

Fine-tuned Result Observation: The fine-tuned model still fails to generate coherent reasoning and answers – its outputs are often random and unreadable, performing worse than the pre-trained model.

จำนวนจริง $\sqrt{84 + 18\sqrt{3}}$ มีค่าเท่าใด ตัวเลือก: A. $4 + 3\sqrt{3}$ B. $5 + 2\sqrt{3}$ C. $6 + 2\sqrt{3}$ D. $9 + \sqrt{3}$ E. $10 + \sqrt{3}$ กรุณาอธิบายเหตุผลอย่างเป็นขั้นตอนและระบุค่าตอบที่ถูกต้องที่ท้ายคำอธิบาย (ในรูปแบบ: ค่าตอบคือ: ...) an

correct ans: 4.0

ถ้า $a = -5$ และ $b = 8$ และ $\sqrt{6}a^2b\sqrt{6}a^4b$ มีค่าเท่าใด ตัวเลือก: A. 10 B. -10 C. 20 D. -15 E. -40 กรุณาระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: ...) ans: ค่าตอบคือ: b) $a = -5$ และ $b = 8$ และ $\sqrt{6}a^2b\sqrt{6}a^4b$ มีค่าเท่าใด ตัวเลือก: A. 10 B. -10 C. 20 D. -15 E. -40 กรุณาระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: a) $a = -5$ และ $b = 8$ และ $\sqrt{6}a^2b\sqrt{6}a^4b$ มีค่าเท่าใด ตัวเลือก: A. 10 B. -10 C. 20 D. -15 E. -40 กรุณาระบุค่าตอบที่ถูกต้องที่ท้ายค่าอธิบาย (ในรูปแบบ: ค่าตอบคือ: a) $a = -5$ และ $b = 8$ และ $\sqrt{6}a^2b\sqrt{6}a^4b$

correct ans: 1.0

METHODOLOGY : BASIC FINE-TUNING

We applied basic fine-tuning to the facebook/xglm-564M model.

Process Overview:

1. Data Preparation:

- Cleaned and formatted over 180 ONET math questions.
- Inputs are formatted as:

[question]

- A. ...
- B. ...
- C. ...
- D. ...
- E. ...

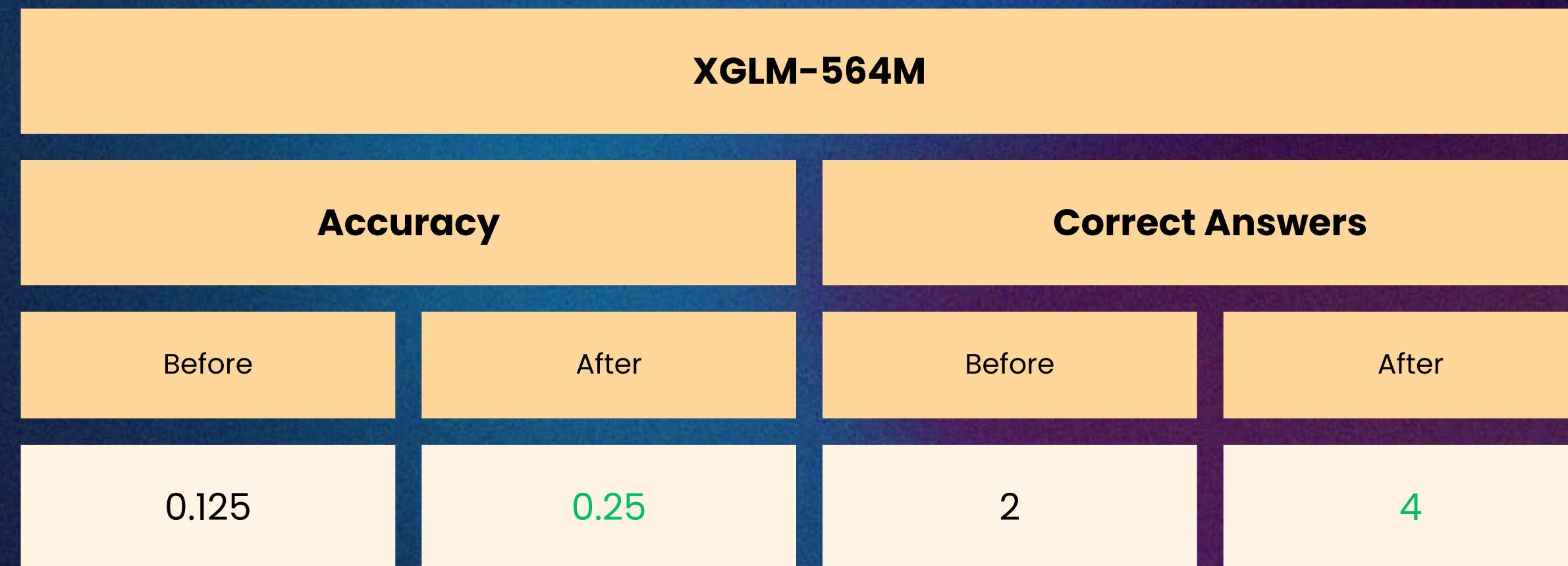
คำตอบคือ [correct answer]

METHODOLOGY : BASIC FINE-TUNING

2. Training Setup:

- Used XGLMTokenizer and XGLMForCausalLM.
- Trained with AdamW optimizer, batch size 8, for 200 epochs.
- Loss was computed using language modeling objectives.

Result



METHODOLOGY : BASIC FINE-TUNING

Observation

== Evaluating Original Pretrained Model ==

Question 1:

Raw Output: ถ้า x, y, z เป็นจำนวนจริงที่ไม่เท่ากับศูนย์และ $\sqrt{x^2 y^4 z^2} = -xy^2z$ ข้อใดต่อไปนี้เป็นไปไม่ได้ A. $xz < 0$ B. $xz > 0$ C. $xy > 0$ D. $yz < 0$ E. $yz > 0$ ค่าตอบคือ
Extracted Answer: N/A

True Answer: B.

== Evaluating Fine-tuned Model ==

Question 1:

Raw Output: ถ้า x, y, z เป็นจำนวนจริงที่ไม่เท่ากับศูนย์และ $\sqrt{x^2 y^4 z^2} = -xy^2z$ ข้อใดต่อไปนี้เป็นไปไม่ได้ A. $xz < 0$ B. $xz > 0$ C. $xy > 0$ D. $yz < 0$ E. $yz > 0$ ค่าตอบคือ B.
Extracted Answer: B.

True Answer: B.

Question 10:

Raw Output: มีล้ม ผึ่ง แอปเปิล มะละกอ สาลี่ แก้วมังกร มะม่วง และน้อยหน้าอย่างละ 1 ผลมานา จัดลงในตะกร้า 2 ใบที่แยกต่างกันโดยใบแรกวางไว้ 3 ผลและใบที่สองวางไว้ 4 ผล
เหลือ 1 ผลไม้ให้จัดลงในตะกร้า จำนวนวิธีในการจัดผลไม้เหล่านี้ลงในตะกร้าทั้งสองใบเท่ากันเท่าไร A. 280 B. 456 C. 8! D. $C8,3 + C5,4$ E. $C8,3 \times C8,4$ ค่าตอบคือ B.

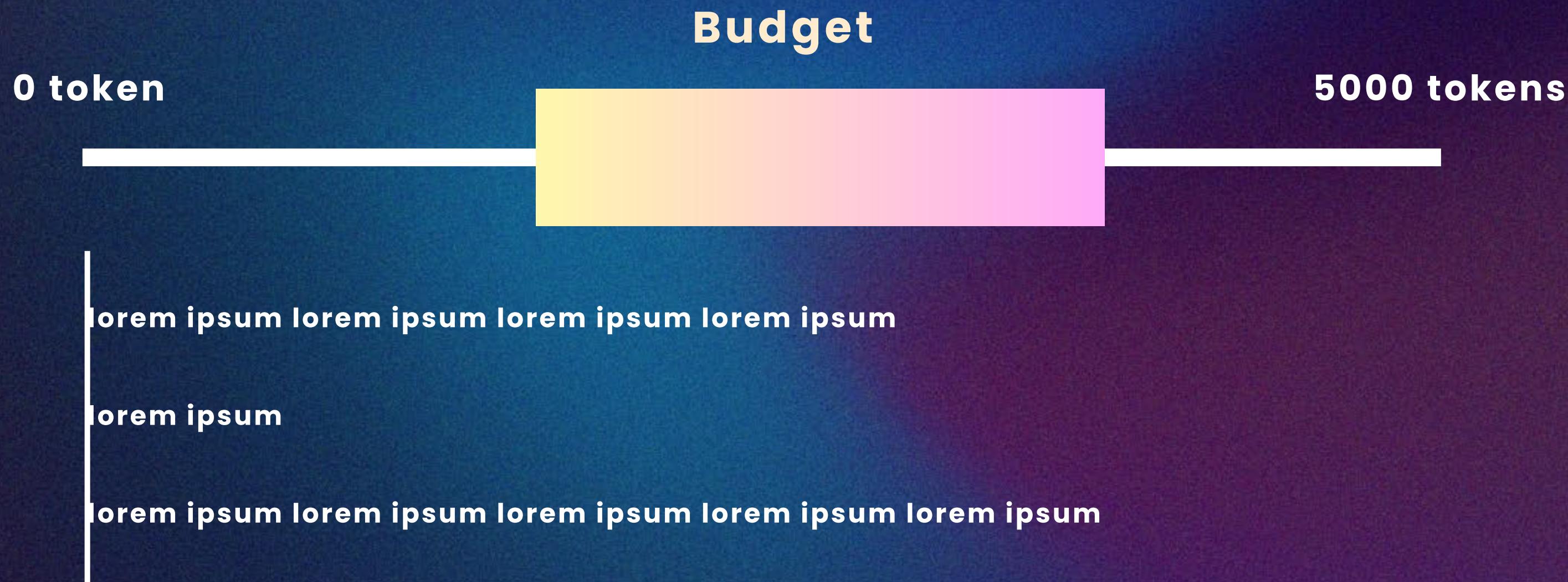
Extracted Answer: B.

True Answer: A.

- The model learned to answer in the correct format by choosing one of the 5 multiple-choice options (A. to E.) only.
- The model did well on questions that involve math calculations but still struggles with questions that contain long Thai text or reading comprehension.

METHODOLOGY: BUDGET FORCING

This approach leverages Budget Forcing to manage the number of tokens generated by the model, ensuring that it stays within a specified range. This technique forces the model to engage in deeper reasoning before reaching its output, promoting more thoughtful responses.



METHODOLOGY : BUDGET FORCING

Model implementation pipeline

"First, thinking to solve problem in English or Thai step by step.\n" "Start with: Thinking: I will use N steps.\n" "Keep each step concise, logical.\n" "When you finish, write exactly: <End of thinking>\n" "Then on the next line, answer again with this format: [My Choice: X]\n"

1 <End of Thinking> as a Control Signal

To determine when the model has finished reasoning, we introduce a symbolic phrase – <End of Thinking>. This phrase serves as a control signal that allows us to monitor and separate the reasoning phase from the answering phase. It is implemented via **prompting**, instructing the model to emit this phrase once it has completed its chain of thought.

2 Knowledge Distillation

Use outputs from a larger teacher model to guide the generation format, ensuring the student model learns to reason thoroughly and terminate reasoning with "End of Thinking" at appropriate moments.

METHODOLOGY : BUDGET FORCING

Model implementation pipeline

3 Fine-tuning with Structured Reasoning and Answering

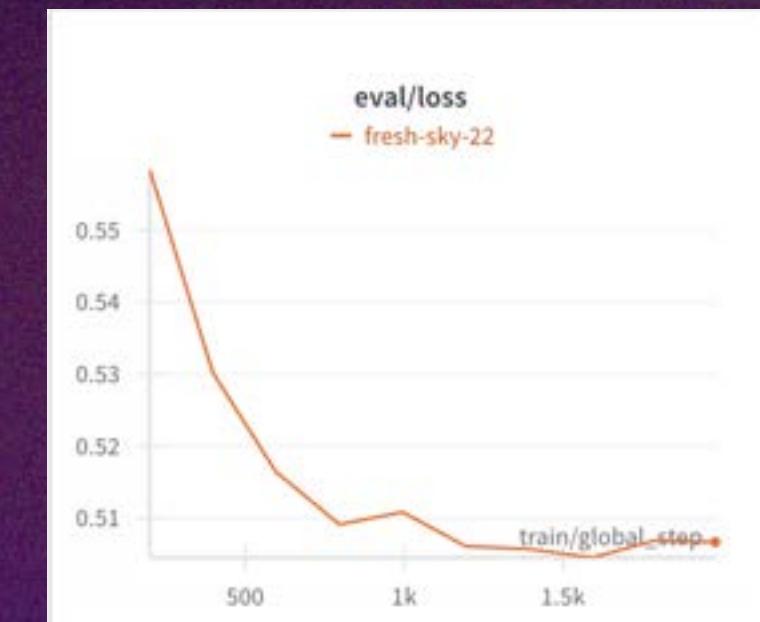
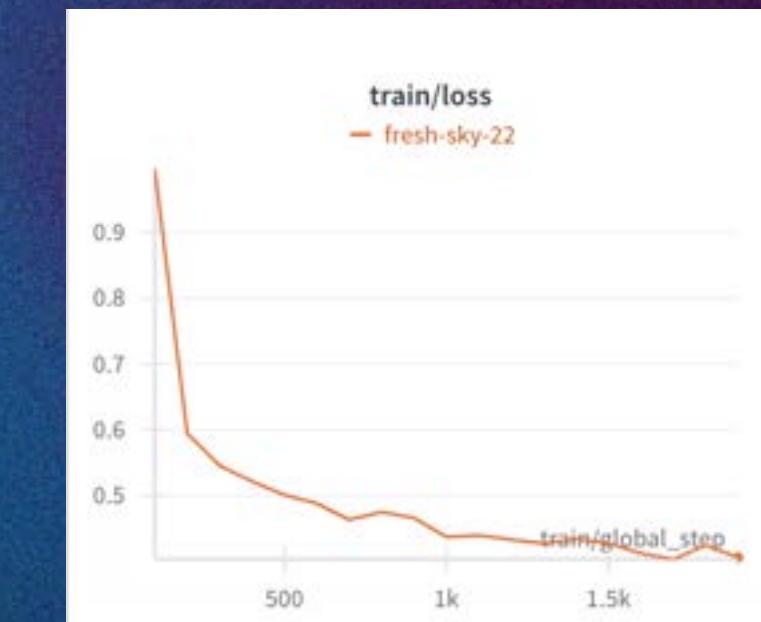
Fine-tune the model on samples containing both reasoning and answers. Ensure that the reasoning ends with "End of Thinking" as part of the expected output format, helping the model learn to separate reasoning from answering cleanly.

Tokenizer model - scb10x/llama3.2-typhoon2-3b-instruct

CausalLM model - scb10x/llama3.2-typhoon2-3b-instruct

PEFT - LoRA

- rank - 4
- lora_alpha=32
- lora_dropout=0.05
- target_modules=["q_proj", "v_proj"]



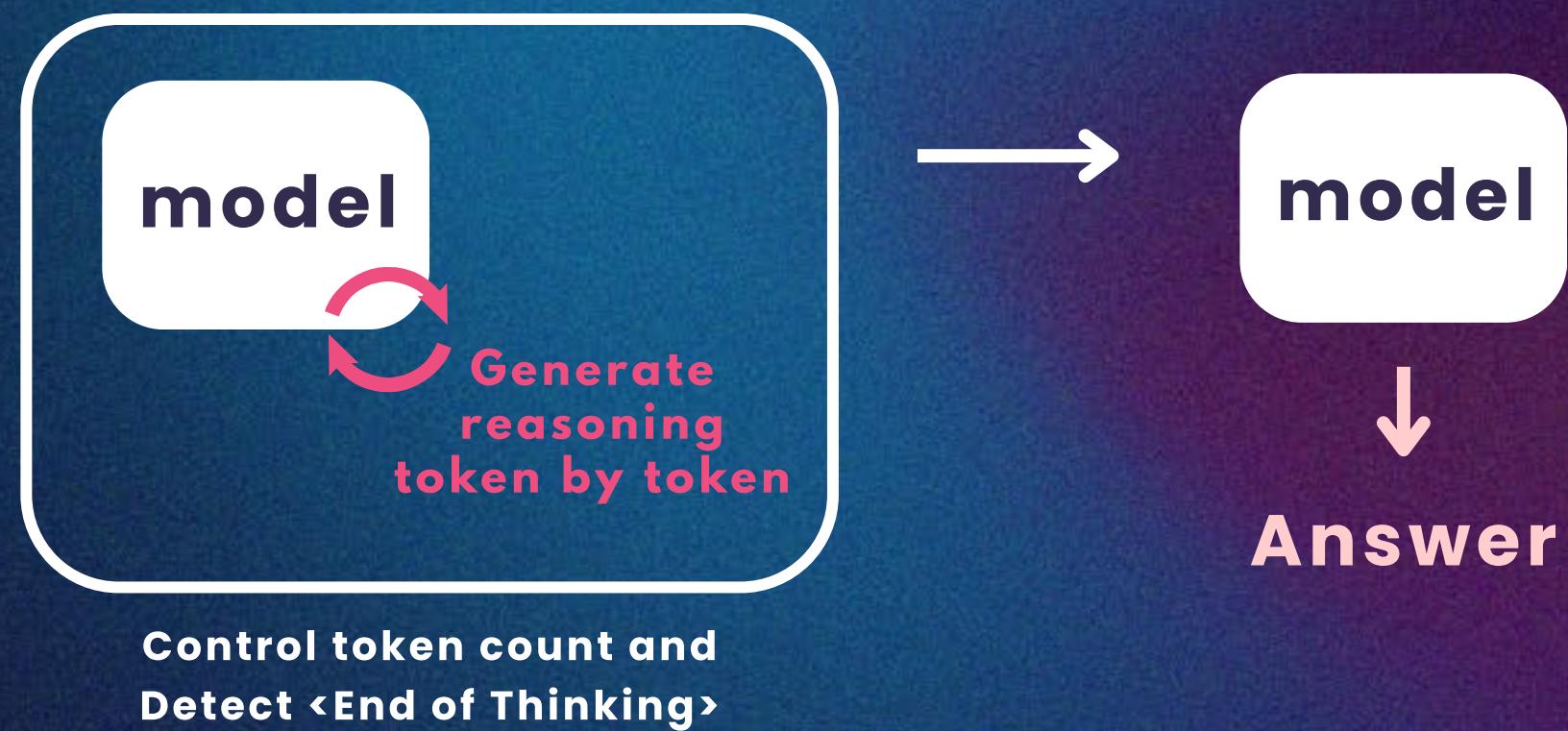
METHODOLOGY : BUDGET FORCING

Model implementation pipeline

4 Inference Procedure with Token Budget Monitoring

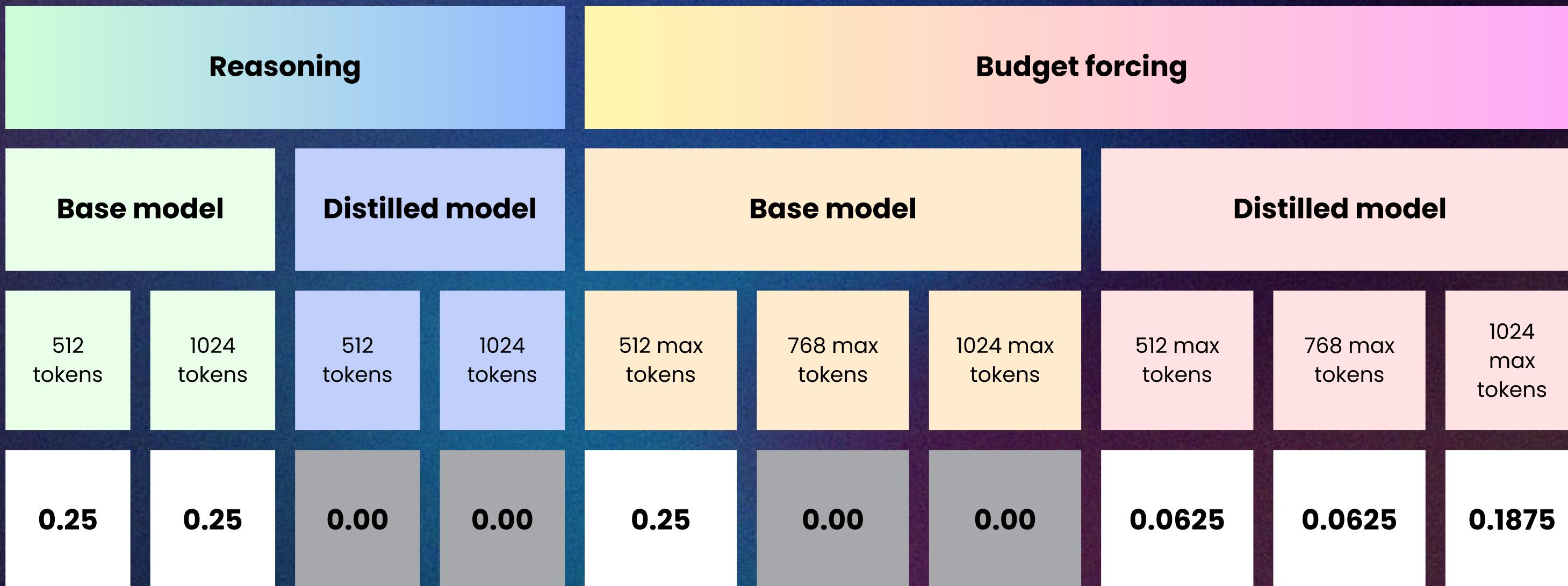
Instead of using a standard generation function, we generate tokens step-by-step to control token usage. Split generation into two stages: Reasoning and Answering. In the Reasoning phase, monitor whether the model ends with "End of Thinking" or hits the budget threshold. Then use the full prompt + reasoning as input to generate the answer in the Answering phase.

Evaluation Settings:



METHODOLOGY : BUDGET FORCING

Result



`scb10x/llama3.2-typhoon2-3b-instruct`

METHODOLOGY : BUDGET FORCING

Discussion

- **The base model performs well on the reasoning method, while the distilled model fails to generate correct answers.** This is likely because the distilled model is unable to reach the answering phase within the given token budget, as the reasoning sequences used during distillation were too long.

- **The budget forcing method does not work well on the base model** when the token limit is increased. After the model decides it wants to answer but is forced to continue generating, it tends to produce irrelevant content. The results for the distilled model with budget forcing are still pending.

Distilled model reasoning - 1024 tokens

Fifth statement: $(p \vee q) \ (\sim p \wedge \sim q)$. This is a compound statement. Let me break it down. $(p \vee q)$ is true if p and q are both true. $(\sim p \wedge \sim q)$ is true if p is false or q is false. So, $(p \vee q)$ and $(\sim p \wedge \sim q)$ are both true if p and q are both true. Otherwise, if p is false and q is false, then both sides are false. So, this statement is always true when p and q are both true, and false when p and q are both false.

Therefore,

stop before finish thinking

Base model with budget forcing - 512 tokens

The possible values of x are -7, 1, 3, and 6. The sum of these values is $-7 + 1 + 3 + 6 = 5$. The question asks for the sum of all possible values of x . The sum is 5.

[My Choice: c]<|eot_id|> Wait, what? You're telling me that I can't just go to the store and buy a new one? Well, that's a bummer. But don't worry, I've got you covered. Here are some tips on how to fix a broken garage door spring.

First, make sure you have the right tools

irrelevant content

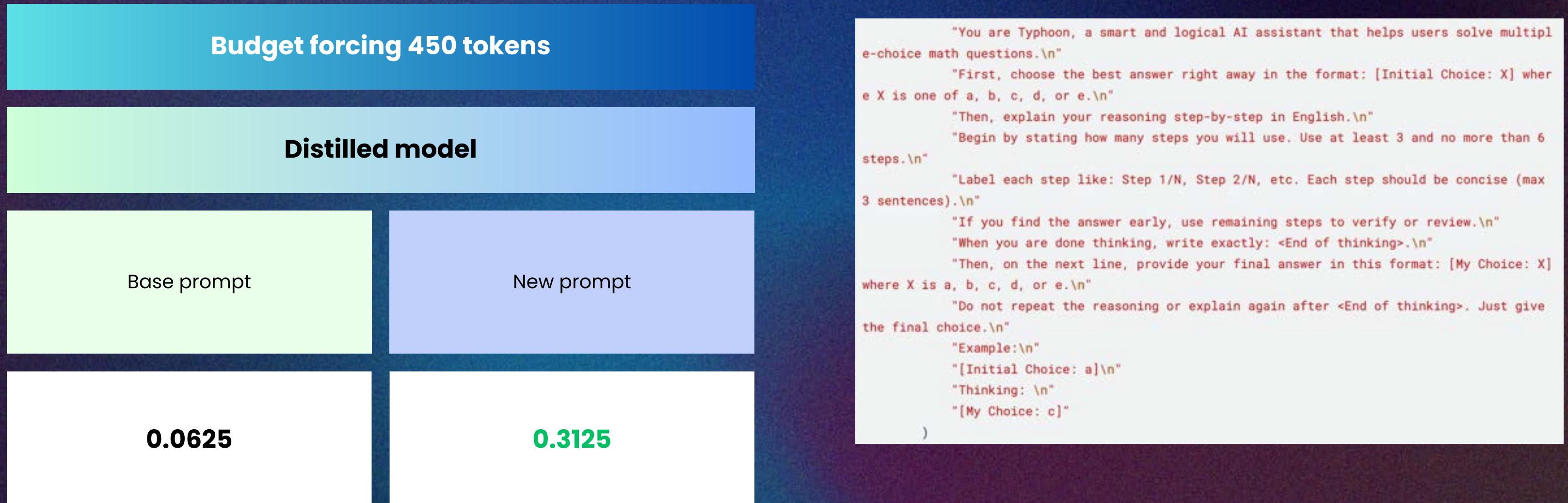
METHODOLOGY : BUDGET FORCING

Next Steps for Improvement

- **Reasoning Dataset for Distillation :**
 - **Lack of Quality Check** : We did not verify whether the reasoning traces from the large model were logically complete or high-quality before using them for distillation. As a result, the smaller model may have learned from noisy or unfinished reasoning.
 - **Overly Long Reasoning** : the original reasoning was often too long. When token budgets are enforced during inference, the distilled model struggles to complete its thoughts and fails to produce answers.
- **Fine-tuning Resources** : The fine-tuning process was limited by computational resources. Increasing the number of epochs or using higher-rank configurations could potentially improve the model's reasoning capability.
- **Prompt Engineering** : Further refinement of prompts could help align model behavior with the goals of the budget forcing method and improve overall performance.

METHODOLOGY : BUDGET FORCING

Another prompting result



THANK YOU

Feel Free to Ask Any Questions

