

ИНТЕЛЛЕКТ ➔

MoE, DeepSeek, Qwen3

к.ф.-м.н. Тихомиров Михаил Михайлович

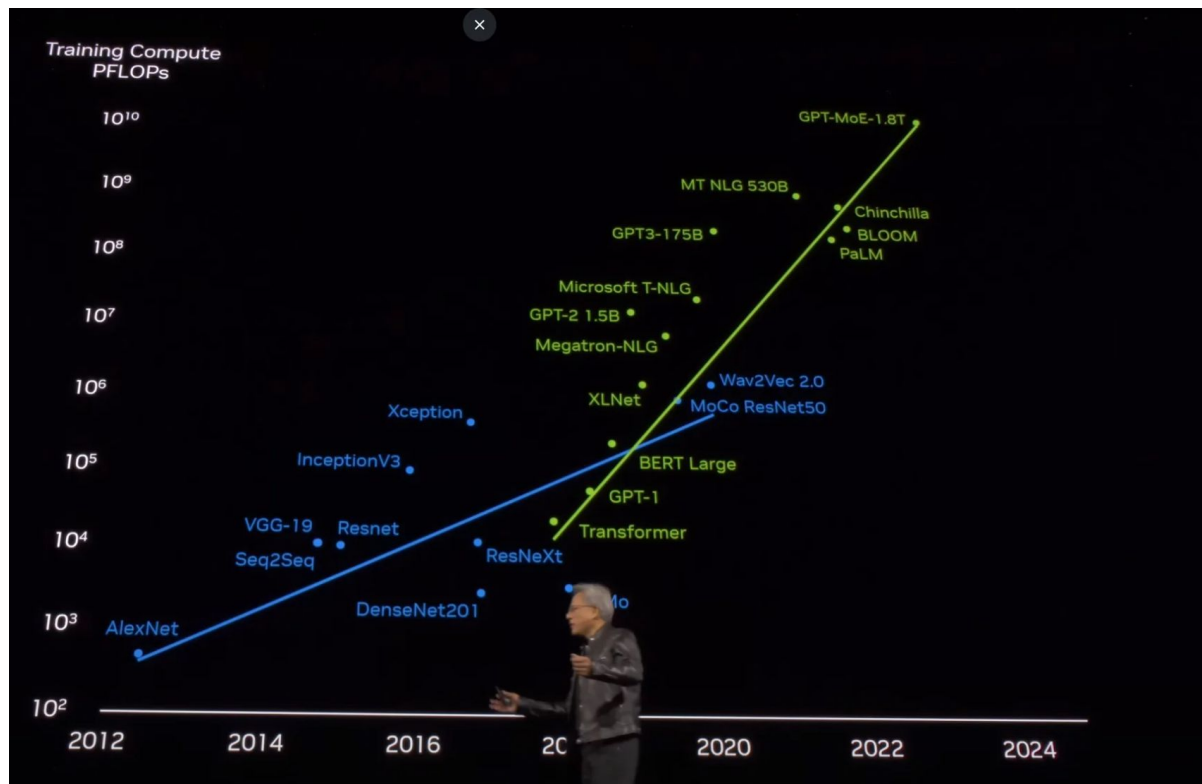
научный сотрудник
НИВЦ МГУ имени М. В. Ломоносова

Mixture of Experts (MoE)

GPT-4 - MoE?



- Доклад NVIDIA на GTC.
- Соответствует ранним слухам про: GPT-4 - 1.8T MoE model с 16 экспертами, 111B каждый



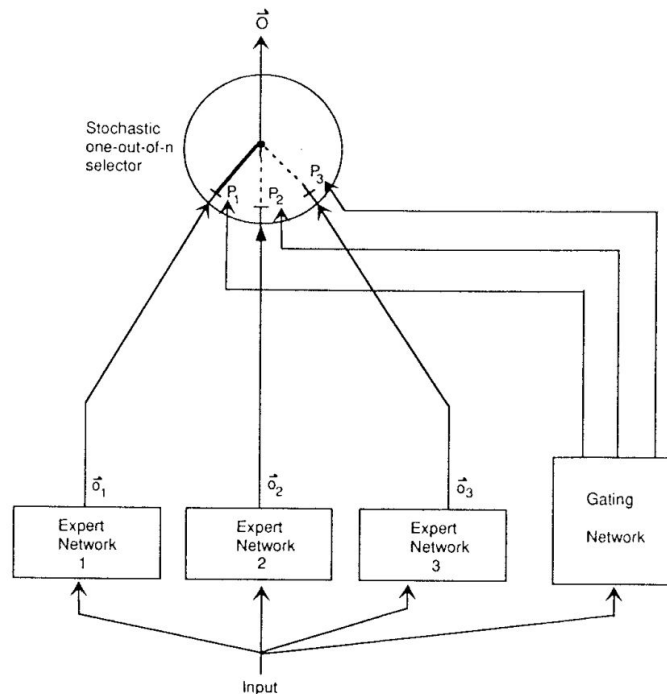
Mixture of Experts (MoE)



- Истоки подхода в 1991 году.
- Одинаковый input идет в 3 FF слоя-эксперта и в роутер “Gating Network”.
- Каждый эксперт должен был отвечать за разные подмножества тренировочной выборки, специализируясь на них.

Adaptive Mixtures of Local Experts

81



MoE: Sparsity

- LSTM с экспертами.
- Идея в Conditional computation.
- Каждый “эксперт” - FeedForward.
- Gating Network обучается и отвечает за выбор экспертов.
- Выбирается только top-k экспертов каждый токен.

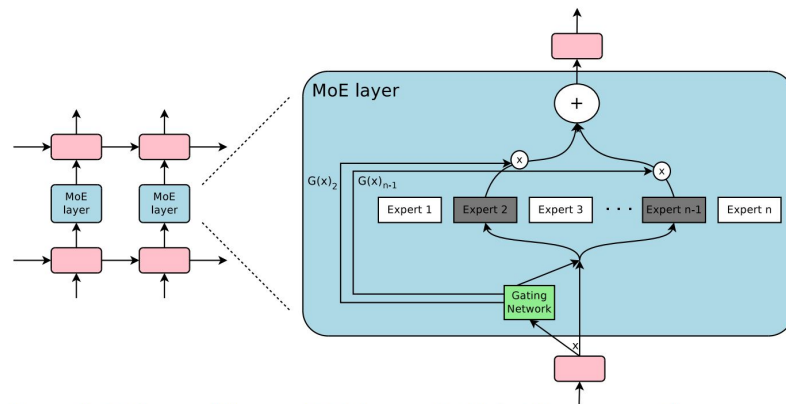


Figure 1: A Mixture of Experts (MoE) layer embedded within a recurrent language model. In this case, the sparse gating function selects two experts to perform computations. Their outputs are modulated by the outputs of the gating network.



MoE: Sparsity

$$x \rightarrow y: \quad y = \sum_{i=1}^n G(x)_i E_i(x)$$

Классический (1994) роутинг:

$$G_{\sigma}(x) = \text{Softmax}(x \cdot W_g)$$

Conditional роутинг (плюс шум):

$$G(x) = \text{Softmax}(\text{KeepTopK}(H(x), k))$$

$$H(x)_i = (x \cdot W_g)_i + \text{StandardNormal}() \cdot \text{Softplus}((x \cdot W_{noise})_i)$$

$$\text{KeepTopK}(v, k)_i = \begin{cases} v_i & \text{if } v_i \text{ is in the top } k \text{ elements of } v. \\ -\infty & \text{otherwise.} \end{cases}$$

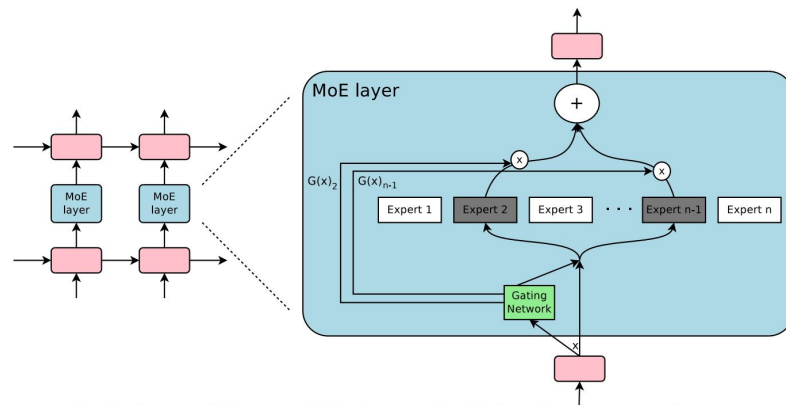


Figure 1: A Mixture of Experts (MoE) layer embedded within a recurrent language model. In this case, the sparse gating function selects two experts to perform computations. Their outputs are modulated by the outputs of the gating network.

LSTM MoE: Results



Table 2: Results on WMT'14 En→Fr newstest2014 (bold values represent best results).

Model	Test Perplexity	Test BLEU	ops/timestep	Total #Parameters	Training Time
MoE with 2048 Experts	2.69	40.35	85M	8.7B	3 days/64 k40s
MoE with 2048 Experts (longer training)	2.63	40.56	85M	8.7B	6 days/64 k40s
GNMT (Wu et al., 2016)	2.79	39.22	214M	278M	6 days/96 k80s
GNMT+RL (Wu et al., 2016)	2.96	39.92	214M	278M	6 days/96 k80s
PBMT (Durrani et al., 2014)		37.0			
LSTM (6-layer) (Luong et al., 2015b)		31.5			
LSTM (6-layer+PosUnk) (Luong et al., 2015b)		33.1			
DeepAtt (Zhou et al., 2016)		37.7			
DeepAtt+PosUnk (Zhou et al., 2016)		39.2			

Table 3: Results on WMT'14 En → De newstest2014 (bold values represent best results).

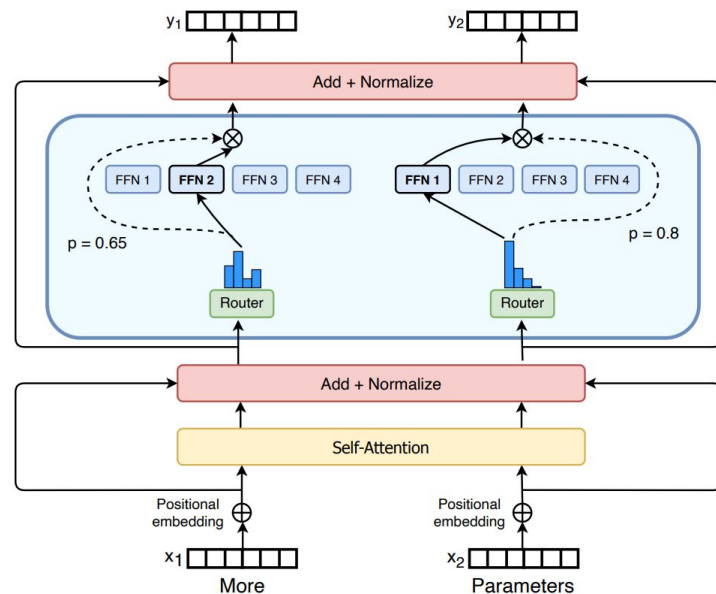
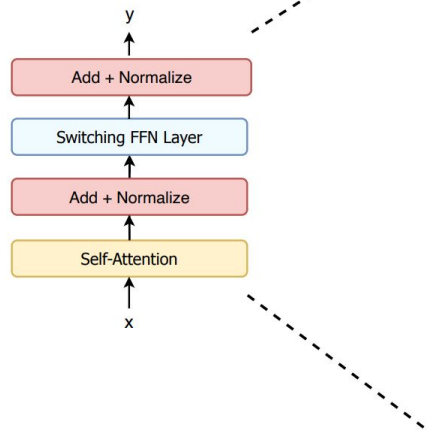
Model	Test Perplexity	Test BLEU	ops/timestep	Total #Parameters	Training Time
MoE with 2048 Experts	4.64	26.03	85M	8.7B	1 day/64 k40s
GNMT (Wu et al., 2016)	5.25	24.91	214M	278M	1 day/96 k80s
GNMT+RL (Wu et al., 2016)	8.08	24.66	214M	278M	1 day/96 k80s
PBMT (Durrani et al., 2014)		20.7			
DeepAtt (Zhou et al., 2016)		20.6			

Table 4: Results on the Google Production En→Fr dataset (bold values represent best results).

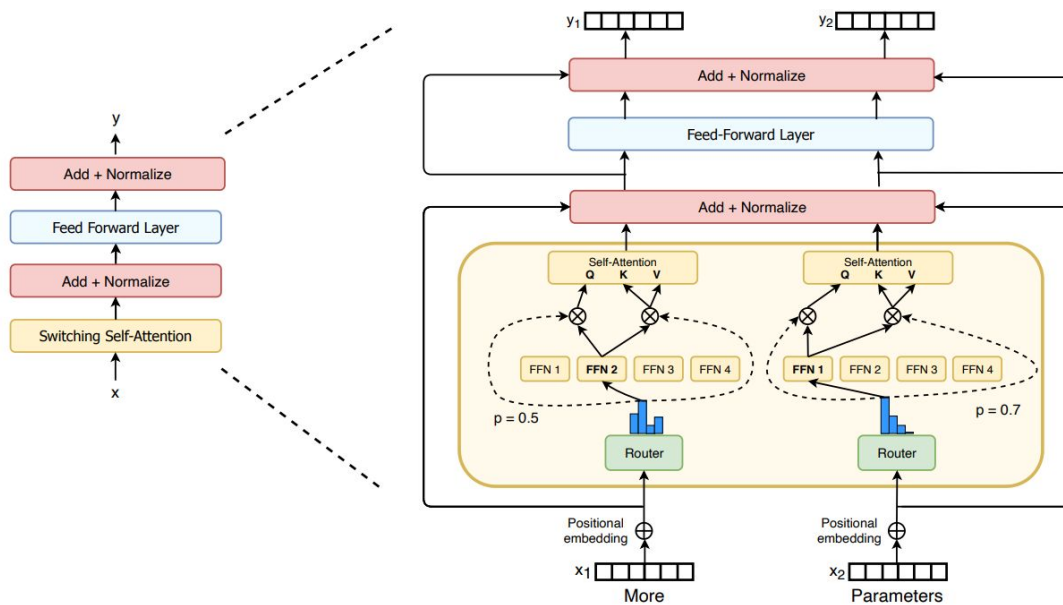
Model	Eval Perplexity	Eval BLEU	Test Perplexity	Test BLEU	ops/timestep	Total #Parameters	Training Time
MoE with 2048 Experts	2.60	37.27	2.69	36.57	85M	8.7B	1 day/64 k40s
GNMT (Wu et al., 2016)	2.78	35.80	2.87	35.56	214M	278M	6 days/96 k80s

Switch Transformer

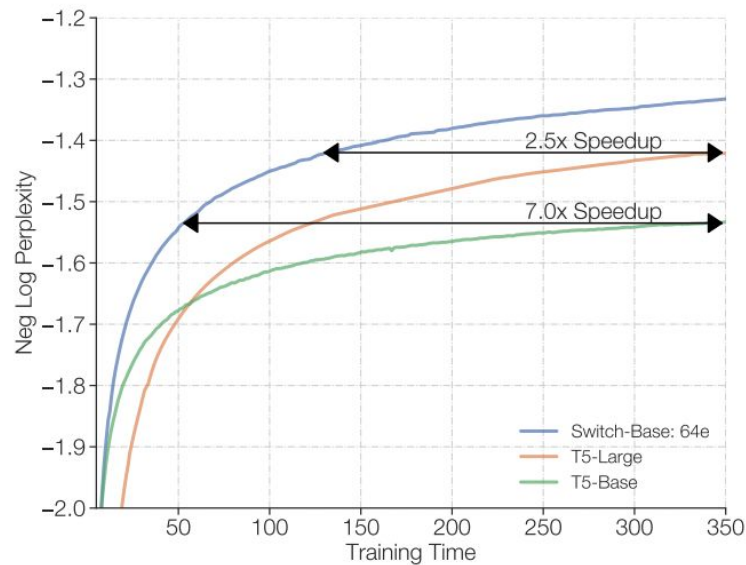
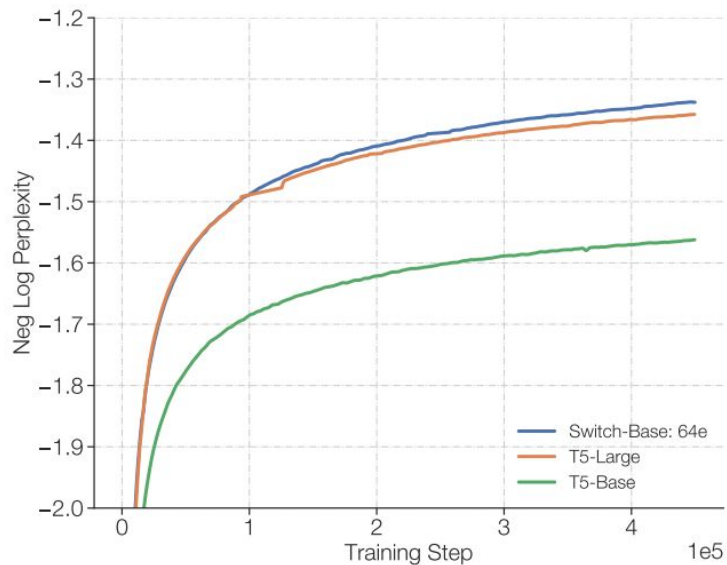
- T5 (encoder-decoder)
- Выбирается 1 эксперт.
- Вспомогательный loss для каждого Switch слоя.



Switch Transformer: Switching Attention



Switch Transformer: ускорение сходимости



Switch Transformer: что лучше роутить?



Model	Precision	Quality @100k Steps (↑)	Quality @16H (↑)	Speed (ex/sec) (↑)
Experts FF	float32	-1.548	-1.614	1480
Expert Attention	float32	-1.524	-1.606	1330
Expert Attention	bfloat16	[diverges]	[diverges]	–
Experts FF + Attention	float32	-1.513	-1.607	1240
Expert FF + Attention	bfloat16	[diverges]	[diverges]	–

Что выучивают эксперты?

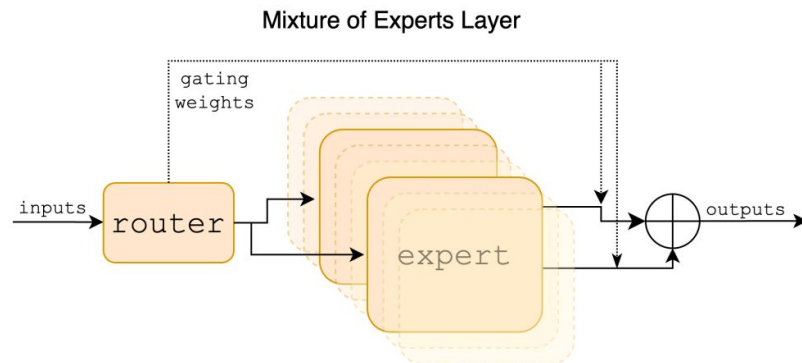


Expert specialization	Expert position	Routed tokens
Sentinel tokens	Layer 1	been <extra_id.4><extra_id.7>floral to <extra_id.10><extra_id.12><extra_id.15> <extra_id.17><extra_id.18><extra_id.19>...
	Layer 4	<extra_id.0><extra_id.1><extra_id.2> <extra_id.4><extra_id.6><extra_id.7> <extra_id.12><extra_id.13><extra_id.14>...
	Layer 6	<extra_id.0><extra_id.4><extra_id.5> <extra_id.6><extra_id.7><extra_id.14> <extra_id.16><extra_id.17><extra_id.18>...
Punctuation	Layer 2	, , , , , , , - , , , , ,))
	Layer 6	, , , , , : : , & , & & ? & - , ? , , , . <extra_id.27>
Conjunctions and articles	Layer 3	The the the the the the the the The the the
	Layer 6	the the the The the the the a and and and and and and or and a and . the the if ? a designed does been is not
Verbs	Layer 1	died falling identified fell closed left posted lost felt left said read miss place struggling falling signed died falling designed based disagree submitted develop
Visual descriptions <i>color, spatial position</i>	Layer 0	her over her know dark upper dark outer center upper blue inner yellow raw mama bright bright over open your dark blue
Proper names	Layer 1	A Mart Gr Mart Kent Med Cor Tri Ca Mart R Mart Lorraine Colin Ken Sam Ken Gr Angel A Dou Now Ga GT Q Ga C Ko C Ko Ga G
Counting and numbers <i>written and numerical forms</i>	Layer 1	after 37 19. 6. 27 I I Seven 25 4, 54 I two dead we Some 2012 who we few lower each

Mixtral 8x7B



- 8 экспертов по 7b (в сумме).
- 2 активных за раз ($K=2$).
 - Итого скорость на инференсе ~как у 14b моделей.
- Длина последовательности: 32k.
- Мультиязычная.



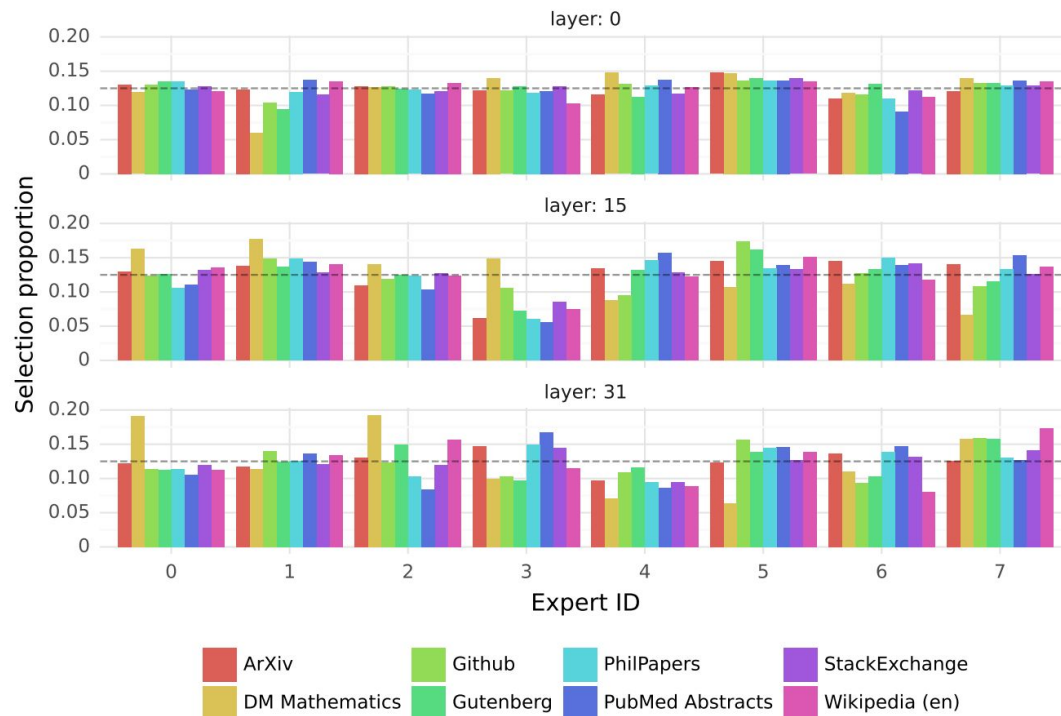
Mixtral 8x7B: результаты



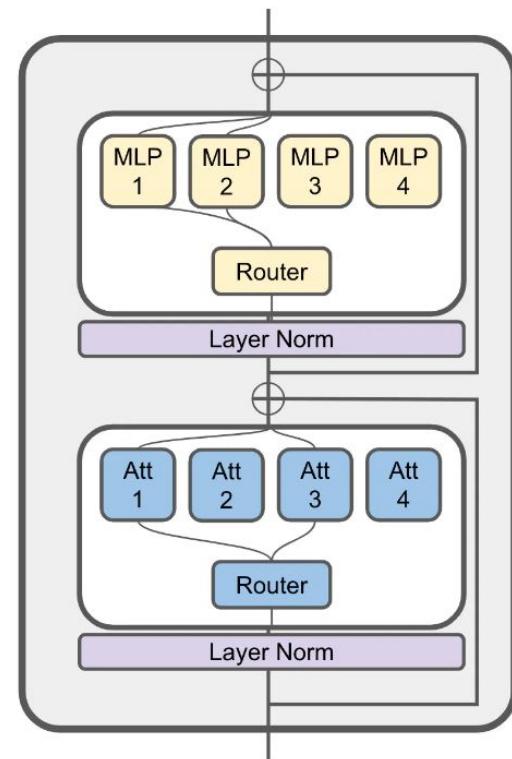
	LLaMA 2 70B	GPT-3.5	Mixtral 8x7B
MMLU (MCQ in 57 subjects)	69.9%	70.0%	70.6%
HellaSwag (10-shot)	87.1%	85.5%	86.7%
ARC Challenge (25-shot)	85.1%	85.2%	85.8%
WinoGrande (5-shot)	83.2%	81.6%	81.2%
MBPP (pass@1)	49.8%	52.2%	60.7%
GSM-8K (5-shot)	53.6%	57.1%	58.4%
MT Bench (for Instruct Models)	6.86	8.32	8.30

Mixtral 8x7B: специализация экспертов

Специализация экспертов по доменам не была обнаружена.



- Комбинация из Mixture of MLP Experts и Mixture of Attention heads.
- 24 блока, 8 экспертов, 2 активных.
- 8 млрд параметров в сумме, ~2 млрд активных во время инференса.
- Low Budget: 0.1M долларов, 96 H100 GPU cluster в течение 2 недель.
- 1.3 триллиона высококачественных ТОКЕНОВ.



JetMoE: результаты



Model	Active Params	Training Tokens	MBPP	Open LLM Leaderboard Average	ARC	Hellaswag	MMLU	TruthfulQA	WinoGrande	GSM 8K
Gemma-2B	2B	2T	28.0	46.4	48.4	71.8	41.8	33.1	66.3	16.9
DeepseekMoE-16B	2.8B	2T	34.0	51.1	53.2	79.8	46.3	36.1	73.7	17.3
LLaMA2-7B	7B	2T	20.8	51.0	53.1	78.6	46.9	38.8	74.0	14.5
LLaMA-13B	13B	1T	22.0	51.4	56.2	80.9	47.7	39.5	76.2	7.6
JetMoE-8B	2.2B	1.25T	34.2	53.0	48.7	80.5	49.2	41.7	70.2	27.8



Mixture of Experts: Выводы

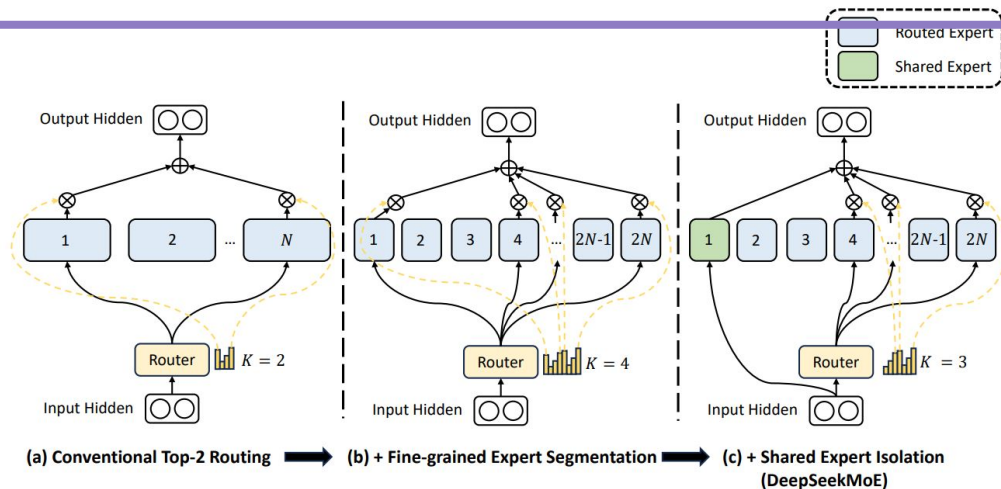
- Идейно подход существует давно и показывал себя достаточно успешно на разных этапах развития отрасли.
- Позволяет ускорить как обучение так и инференс по сравнению с классическими моделями.
- Потребляет **существенно** большее количество видео памяти.
- Есть проблемы с fine-tuning итоговой модели, но последние работы их вроде как решают.
- Nvidia анонсировала новые видеокарты, которые имеют существенно больший размер видео памяти (B200 192GB)

DeepSeek

DeepSeek: эксперты



- Fine-grained Experts: разрезаем наших экспертов на M меньших, роутинг на уровне меньших экспертов.
- Shared Experts - несколько экспертов активны всегда

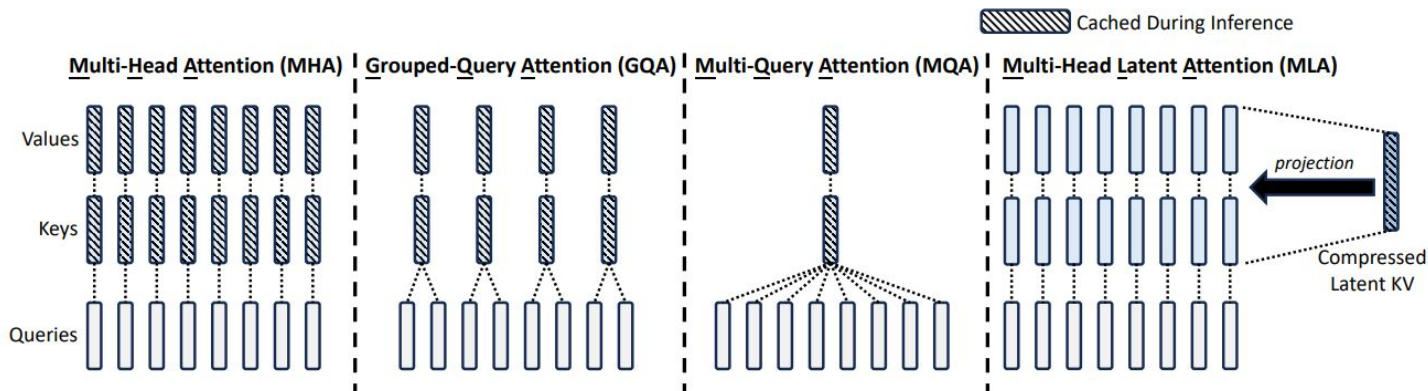


$$\mathbf{h}_t^l = \sum_{i=1}^{K_s} \text{FFN}_i(\mathbf{u}_t^l) + \sum_{i=K_s+1}^{mN} (g_{i,t} \text{FFN}_i(\mathbf{u}_t^l)) + \mathbf{u}_t^l,$$

$$g_{i,t} = \begin{cases} s_{i,t}, & s_{i,t} \in \text{Topk}(\{s_{j,t} | K_s + 1 \leq j \leq mN\}, mK - K_s), \\ 0, & \text{otherwise,} \end{cases}$$

$$s_{i,t} = \text{Softmax}_i(\mathbf{u}_t^{lT} \mathbf{e}_i^l).$$

DeepSeek: MLA



- Вместо получение \mathbf{k} и \mathbf{v} из \mathbf{h} напрямую возникает промежуточное звено \mathbf{c} .
Кэшируем только \mathbf{c} !
- Экономия памяти!

$$\mathbf{c}_t^{KV} = W^{DKV} \mathbf{h}_t,$$

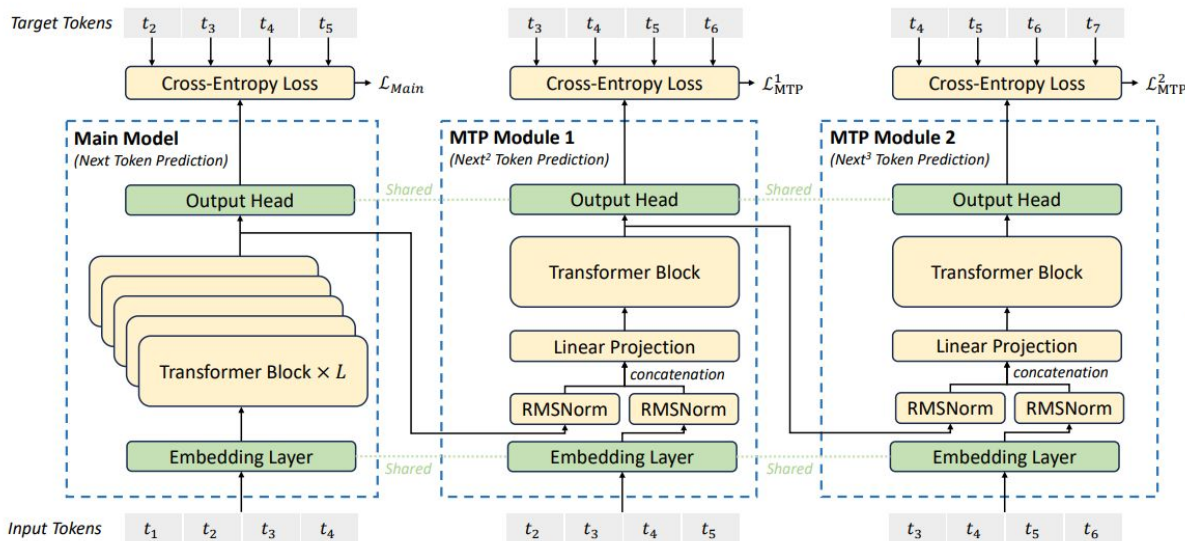
$$\mathbf{k}_t^C = W^{UK} \mathbf{c}_t^{KV},$$

$$\mathbf{v}_t^C = W^{UV} \mathbf{c}_t^{KV},$$

$$\mathbf{c}_t^Q = W^{DQ} \mathbf{h}_t,$$

$$\mathbf{q}_t^C = W^{UQ} \mathbf{c}_t^Q,$$

DeepSeek: MTP



$$\mathcal{L}_{MTP}^k = \text{CrossEntropy}(P_{2+k:T+1}^k, t_{2+k:T+1}) = -\frac{1}{T} \sum_{i=2+k}^{T+1} \log P_i^k[t_i],$$

$$\mathcal{L}_{MTP} = \frac{\lambda}{D} \sum_{k=1}^D \mathcal{L}_{MTP}^k.$$

DeepSeek: параметры



Training Costs	Pre-Training	Context Extension	Post-Training	Total
in H800 GPU Hours	2664K	119K	5K	2788K
in USD	\$5.328M	\$0.238M	\$0.01M	\$5.576M

- 671B параметров, 37B активных
- 1 shared и 256 routed экспертов, 8 активных
- Родные веса - FP8, что означает, что нужны H100+ карты
- Для инференса требуется всего 2x8xH100 :)
- В случае A100 -> нужны либо кванты в 4-8 бит, либо bf16 версия (x2 потребность в картах)
- 14.8T в обучении

DeepSeek: результаты

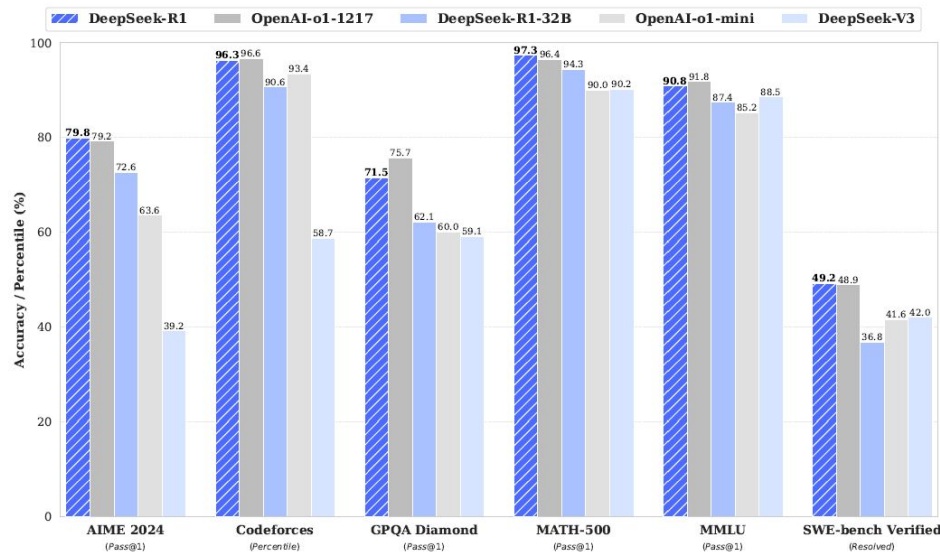


	Benchmark (Metric)	# Shots	DeepSeek-V2 Base	Qwen2.5 72B Base	LLaMA-3.1 405B Base	DeepSeek-V3 Base
	Architecture	-	MoE	Dense	Dense	MoE
	# Activated Params	-	21B	72B	405B	37B
	# Total Params	-	236B	72B	405B	671B
English	Pile-test (BPB)	-	0.606	0.638	0.542	0.548
	BBH (EM)	3-shot	78.8	79.8	82.9	87.5
	MMLU (EM)	5-shot	78.4	85.0	84.4	87.1
	MMLU-Redux (EM)	5-shot	75.6	83.2	81.3	86.2
	MMLU-Pro (EM)	5-shot	51.4	58.3	52.8	64.4
	DROP (F1)	3-shot	80.4	80.6	86.0	89.0
	ARC-Easy (EM)	25-shot	97.6	98.4	98.4	98.9
	ARC-Challenge (EM)	25-shot	92.2	94.5	95.3	95.3
	HellaSwag (EM)	10-shot	87.1	84.8	89.2	88.9
	PIQA (EM)	0-shot	83.9	82.6	85.9	84.7
	WinoGrande (EM)	5-shot	86.3	82.3	85.2	84.9
	RACE-Middle (EM)	5-shot	73.1	68.1	74.2	67.1
	RACE-High (EM)	5-shot	52.6	50.3	56.8	51.3
	TriviaQA (EM)	5-shot	80.0	71.9	82.7	82.9
	NaturalQuestions (EM)	5-shot	38.6	33.2	41.5	40.0
AGIEval (EM)	0-shot	57.5	75.8	60.6	79.6	
Code	HumanEval (Pass@1)	0-shot	43.3	53.0	54.9	65.2
	MBPP (Pass@1)	3-shot	65.0	72.6	68.4	75.4
	LiveCodeBench-Base (Pass@1)	3-shot	11.6	12.9	15.5	19.4
	CRUXEval-I (EM)	2-shot	52.5	59.1	58.5	67.3
	CRUXEval-O (EM)	2-shot	49.8	59.9	59.9	69.8
Math	GSM8K (EM)	8-shot	81.6	88.3	83.5	89.3
	MATH (EM)	4-shot	43.4	54.4	49.0	61.6
	MGSM (EM)	8-shot	63.6	76.2	69.9	79.8
	CMath (EM)	3-shot	78.7	84.5	77.3	90.7
Chinese	CLUEWSC (EM)	5-shot	82.0	82.5	83.0	82.7
	C-Eval (EM)	5-shot	81.4	89.2	72.5	90.1
	CMMLU (EM)	5-shot	84.0	89.5	73.7	88.8
	CMRC (EM)	1-shot	77.4	75.8	76.0	76.3
	C3 (EM)	0-shot	77.4	76.7	79.7	78.6
	CCPM (EM)	0-shot	93.0	88.5	78.6	92.0
Multilingual	MMMLU-non-English (EM)	5-shot	64.0	74.8	73.8	79.4

DeepSeek R1



- Популярность началась именно с R1 (хотя V3 вышла несколько раньше)
- Является одной из первых reasoning открытых моделей!
- RL (метод GRPO) на основе rule-based награды поверх V3
- Выпустили кучу разных distill версий
- Соперник для O1



Reasoning модели



Chatbot

Сколько будет $-1 \times -1 \times -1$?

<think>

Хорошо, давайте разберемся с этим примером: $-1 \times -1 \times -1$. Мне нужно найти результат умножения трех отрицательных единиц. Я помню, что при умножении чисел с одинаковыми знаками результат положительный, а при разных — отрицательный. Но здесь три числа, и все они отрицательные. Возможно, я должен выполнять умножение по порядку, два первых числа сначала, а потом третье.

Давайте начнем с первого шага: умножить первые два числа. -1×-1 . Как я уже думал, если оба числа отрицательны, то их произведение положительное. Значит, $-1 \times -1 = 1$. Теперь у нас остается умножить этот результат на третье число, которое тоже -1 . То есть 1×-1 . Здесь знаки разные: положительное и отрицательное. Тогда результат должен быть отрицательным. $1 \times -1 = -1$, значит, $1 \times -1 = -1$. Получается, что $-1 \times -1 \times -1 = -1$.

Но я хочу убедиться, что не ошибся. Может быть, есть другой способ проверить это. Например, использовать правило знаков для умножения нескольких чисел. Если количество отрицательных множителей четное, то результат положительный; если нечетное — отрицательный. В данном случае у нас три отрицательных множителя, что нечетное число. Значит, результат должен быть отрицательным. А произведение абсолютных значений всех чисел: $1 \times 1 = 1$. Поэтому окончательный ответ -1 . Это совпадает с предыдущим расчетом. Значит, все правильно.

Еще можно представить это геометрически или через движение на числовой прямой. Умножение на -1 меняет знак числа. Первый шаг: $-1 \times -1 = 1$ (переключили знак дважды, вернулись к положительному). Затем умножаем на -1 еще раз: $1 \times -1 = -1$ (переключили знак обратно). Таким образом, после трех перемножений получаем -1 .

Возможно, кто-то может запутаться, считая, что два минуса дают плюс, и тогда третий минус просто остается. Но важно помнить, что каждый дополнительный умножающийся минус изменяет знак результата. Сначала два минуса дают плюс, затем умножение на минус дает минус. Все логично.

Еще пример: если бы было четыре единицы с минусами, то результат был бы положительным, так как четное количество отрицательных множителей. Но в нашем случае три, поэтому отрицательный. Все сходится.

</think>

Результат умножения

$$-1 \times -1 \times -1$$

равен **-1**.

Объяснение:

1. Умножение первых двух отрицательных чисел:

$$(-1) \times (-1) = 1$$

(Поскольку произведение двух отрицательных чисел положительно).

2. Умножение полученного положительного результата на третье отрицательное число:

$$1 \times (-1) = -1$$

(Поскольку произведение числа с противоположными знаками отрицательно).

Итог:

$$(-1) \times (-1) \times (-1) = -1$$

DeepSeek R1 distill версии



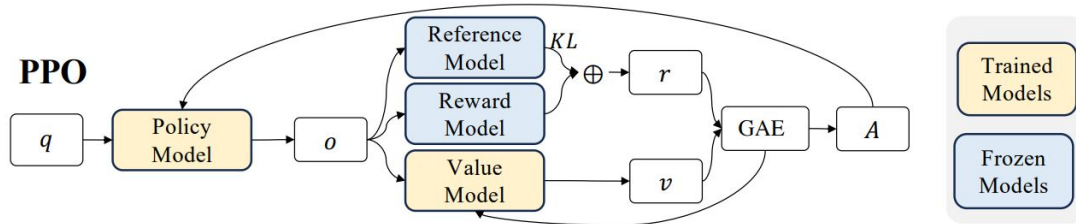
Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	1820
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633

- Дистилл модели были получены путем SFT тюна на генерациях R1
- В их экспериментах Distill из R1 > RL на 32B модели.
- 800т. примеров для дистилляции

GRPO vs PPO



PPO требует некоторую Value model для оценки Advantage на основе текущего Reward.



$$\mathcal{J}_{PPO}(\theta) = \mathbb{E}[q \sim P(Q), o \sim \pi_{\theta_{old}}(O|q)] \frac{1}{|o|} \sum_{i=1}^{|o|} \min \left[\frac{\pi_{\theta}(o_t|q, o_{<t})}{\pi_{\theta_{old}}(o_t|q, o_{<t})} A_t, \text{clip} \left(\frac{\pi_{\theta}(o_t|q, o_{<t})}{\pi_{\theta_{old}}(o_t|q, o_{<t})}, 1 - \epsilon, 1 + \epsilon \right) A_t \right],$$

GRPO - отказ от Value model в сторону генераций N вариантов и расчета средней награды!

$$\mathcal{J}_{GRPO}(\theta) = \mathbb{E}[q \sim P(Q), \{o_i\}_{i=1}^G \sim \pi_{\theta_{old}}(O|q)] \frac{1}{G} \sum_{i=1}^G \left(\min \left(\frac{\pi_{\theta}(o_i|q)}{\pi_{\theta_{old}}(o_i|q)} A_i, \text{clip} \left(\frac{\pi_{\theta}(o_i|q)}{\pi_{\theta_{old}}(o_i|q)}, 1 - \epsilon, 1 + \epsilon \right) A_i \right) - \beta \mathbb{D}_{KL}(\pi_{\theta} || \pi_{ref}) \right), \quad (1)$$

$$\mathbb{D}_{KL}(\pi_{\theta} || \pi_{ref}) = \frac{\pi_{ref}(o_i|q)}{\pi_{\theta}(o_i|q)} - \log \frac{\pi_{ref}(o_i|q)}{\pi_{\theta}(o_i|q)} - 1, \quad (2)$$

where ϵ and β are hyper-parameters, and A_i is the advantage, computed using a group of rewards $\{r_1, r_2, \dots, r_G\}$ corresponding to the outputs within each group:

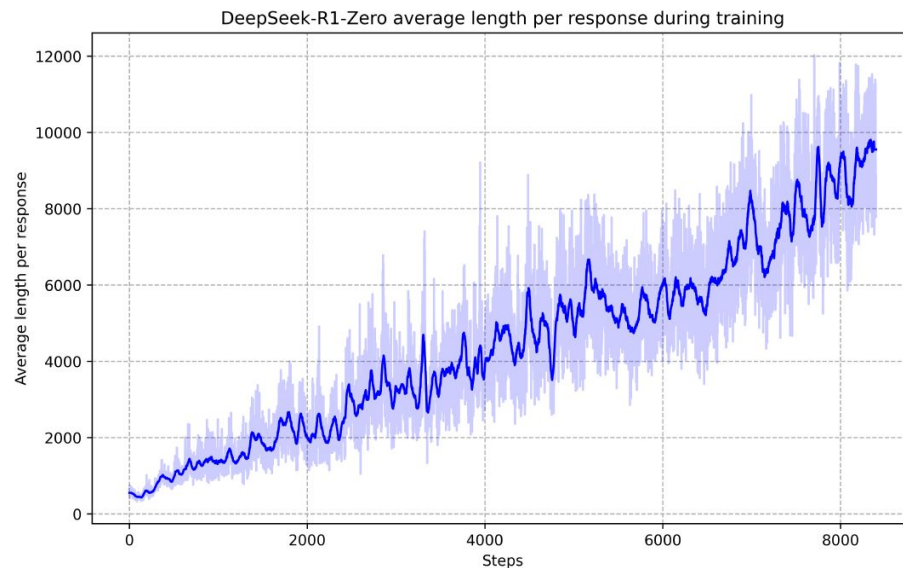
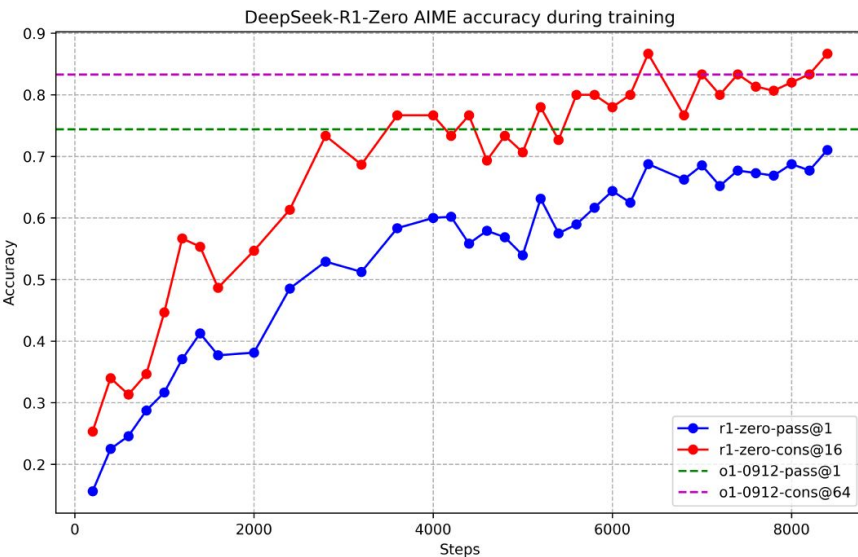
$$A_i = \frac{r_i - \text{mean}(\{r_1, r_2, \dots, r_G\})}{\text{std}(\{r_1, r_2, \dots, r_G\})}. \quad (3)$$

GRPO верифицируемые награды



- **Rule based** - чтобы модель не взламывала награду
- **Accuracy rewards**: ответ, представленный в конце должен быть правильным
- **Format rewards**: блок `<think>...</think>`, ответ в `\boxed{}` ...
- Обучали на математике / коде.

GRPO: step vs accuracy vs len



Qwen3

Qwen3 в релизе!



Почему это важно:

- Лучшие открытые модели на текущий момент в своих весовых
- Интересная размерная сетка
 - от 0.6B до 235B-A22B
- Гибридный ризонер
- 128 экспертов, 8 активных



Размерная сетка



Models	Layers	Heads (Q / KV)	Tie Embedding	Context Length
Qwen3-0.6B	28	16 / 8	Yes	32K
Qwen3-1.7B	28	16 / 8	Yes	32K
Qwen3-4B	36	32 / 8	Yes	32K
Qwen3-8B	36	32 / 8	No	128K
Qwen3-14B	40	40 / 8	No	128K
Qwen3-32B	64	64 / 8	No	128K

Models	Layers	Heads (Q / KV)	# Experts (Total / Activated)	Context Length
Qwen3-30B-A3B	48	32 / 4	128 / 8	128K
Qwen3-235B-A22B	94	64 / 4	128 / 8	128K

Qwen3: pre-training



- Шаг 1: обычный pre-training
 - **30 триллионов** токенов, 119 языков
 - 4096 длина последовательности
- Шаг 2: ризонинг!
 - **5 триллионов** высококачественных токенов, содержащих STEM, код, рассуждения, синтетические данные
 - “We also accelerate the learning rate decay during this stage”
- Шаг 3: Длинный контекст
 - Данные содержащие примеры до длиной 32768 токенов
 - **100+ миллиардов** токенов

Претрейнинг: результаты Qwen3-235B-A22B



Table 3: Comparison among Qwen3-235B-A22B-Base and other representative strong open-source baselines. The highest, the second-best scores are shown in bold and underlined, respectively.

	Qwen2.5-72B Base	Qwen2.5-Plus Base	Llama-4-Maverick Base	DeepSeek-V3 Base	Qwen3-235B-A22B Base
Architecture	Dense	MoE	MoE	MoE	MoE
# Total Params	72B	271B	402B	671B	235B
# Activated Params	72B	37B	17B	37B	22B
<i>General Tasks</i>					
MMLU	86.06	85.02	85.16	<u>87.19</u>	87.81
MMLU-Redux	83.91	82.69	84.05	<u>86.14</u>	87.40
MMLU-Pro	58.07	63.52	<u>63.91</u>	59.84	68.18
SuperGPQA	36.20	37.18	<u>40.85</u>	<u>41.53</u>	44.06
BBH	<u>86.30</u>	85.60	83.62	86.22	88.87
<i>Math & STEM Tasks</i>					
GPQA	<u>45.88</u>	41.92	43.94	41.92	47.47
GSM8K	91.50	<u>91.89</u>	87.72	87.57	94.39
MATH	62.12	<u>62.78</u>	<u>63.32</u>	62.62	71.84
<i>Coding Tasks</i>					
EvalPlus	65.93	61.43	<u>68.38</u>	63.75	77.60
MultiPL-E	58.70	62.16	57.28	<u>62.26</u>	65.94
MBPP	<u>76.00</u>	74.60	75.40	<u>74.20</u>	81.40
CRUX-O	66.20	68.50	<u>77.00</u>	76.60	79.00
<i>Multilingual Tasks</i>					
MGSM	82.40	82.21	79.69	<u>82.68</u>	83.53
MMMLU	84.40	83.49	83.09	<u>85.88</u>	86.70
INCLUDE	69.05	66.97	<u>73.47</u>	<u>75.17</u>	73.46

Претрейнинг: результаты Qwen3-32B



Table 4: Comparison among Qwen3-32B-Base and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.

	Qwen2.5-32B Base	Qwen2.5-72B Base	Gemma-3-27B Base	Llama-4-Scout Base	Qwen3-32B Base
Architecture	Dense	Dense	Dense	MoE	Dense
# Total Params	32B	72B	27B	109B	32B
# Activated Params	32B	72B	27B	17B	32B
<i>General Tasks</i>					
MMLU	83.32	86.06	78.69	78.27	<u>83.61</u>
MMLU-Redux	81.97	83.91	76.53	71.09	<u>83.41</u>
MMLU-Pro	55.10	<u>58.07</u>	52.88	56.13	<u>65.54</u>
SuperGPQA	33.55	<u>36.20</u>	29.87	26.51	39.78
BBH	84.48	<u>86.30</u>	79.95	82.40	87.38
<i>Math & STEM Tasks</i>					
GPQA	<u>47.97</u>	45.88	26.26	40.40	49.49
GSM8K	<u>92.87</u>	91.50	81.20	85.37	93.40
MATH	57.70	62.12	51.78	51.66	<u>61.62</u>
<i>Coding Tasks</i>					
EvalPlus	<u>66.25</u>	65.93	55.78	59.90	72.05
MultiPL-E	58.30	<u>58.70</u>	45.03	47.38	67.06
MBPP	73.60	<u>76.00</u>	68.40	68.60	78.20
CRUX-O	<u>67.80</u>	66.20	60.00	61.90	72.50
<i>Multilingual Tasks</i>					
MGSM	78.12	<u>82.40</u>	73.74	79.93	83.06
MMMLU	82.40	84.40	77.62	74.83	<u>83.83</u>
INCLUDE	64.35	69.05	<u>68.94</u>	68.09	<u>67.87</u>

Претрейнинг: результаты Qwen3-14B/30B-A3B



Table 5: Comparison among Qwen3-14B-Base, Qwen3-30B-A3B-Base, and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.

	Gemma-3-12B Base	Qwen2.5-14B Base	Qwen2.5-32B Base	Qwen2.5-Turbo Base	Qwen3-14B Base	Qwen3-30B-A3B Base
Architecture	Dense	Dense	Dense	MoE	Dense	MoE
# Total Params	12B	14B	32B	42B	14B	30B
# Activated Params	12B	14B	32B	6B	14B	3B
<i>General Tasks</i>						
MMLU	73.87	79.66	83.32	79.50	81.05	<u>81.38</u>
MMLU-Redux	70.70	76.64	81.97	77.11	79.88	<u>81.17</u>
MMLU-Pro	44.91	51.16	55.10	55.60	<u>61.03</u>	61.49
SuperGPQA	24.61	30.68	33.55	31.19	<u>34.27</u>	35.72
BBH	74.28	78.18	84.48	76.10	81.07	<u>81.54</u>
<i>Math & STEM Tasks</i>						
GPQA	31.31	32.83	47.97	41.41	39.90	<u>43.94</u>
GSM8K	78.01	90.22	92.87	88.32	<u>92.49</u>	91.81
MATH	44.43	55.64	57.70	55.60	<u>62.02</u>	<u>59.04</u>
<i>Coding Tasks</i>						
EvalPlus	52.65	60.70	66.25	61.23	72.23	<u>71.45</u>
MultiPL-E	43.03	54.79	58.30	53.24	<u>61.69</u>	66.53
MBPP	60.60	69.00	<u>73.60</u>	67.60	73.40	74.40
CRUX-O	52.00	61.10	<u>67.80</u>	60.20	68.60	67.20
<i>Multilingual Tasks</i>						
MGSM	64.35	74.68	78.12	70.45	79.20	<u>79.11</u>
MMMLU	72.50	78.34	82.40	79.76	79.69	<u>81.46</u>
INCLUDE	63.34	60.26	64.35	59.25	<u>64.55</u>	67.00

Претрейнинг: результаты Qwen3-8B



Table 6: Comparison among Qwen8B-Base and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.

	Llama-3-8B Base	Qwen2.5-7B Base	Qwen2.5-14B Base	Qwen3-8B Base
Architecture	Dense	Dense	Dense	Dense
# Total Params	8B	7B	14B	8B
# Activated Params	8B	7B	14B	8B
<i>General Tasks</i>				
MMLU	66.60	74.16	79.66	76.89
MMLU-Redux	61.59	71.06	76.64	<u>76.17</u>
MMLU-Pro	35.36	45.00	<u>51.16</u>	56.73
SuperGPQA	20.54	26.34	<u>30.68</u>	31.64
BBH	57.70	70.40	<u>78.18</u>	78.40
<i>Math & STEM Tasks</i>				
GPQA	25.80	36.36	32.83	44.44
GSM8K	55.30	85.36	90.22	<u>89.84</u>
MATH	20.50	49.80	<u>55.64</u>	60.80
<i>Coding Tasks</i>				
EvalPlus	44.13	<u>62.18</u>	60.70	67.65
MultiPL-E	31.45	50.73	<u>54.79</u>	58.75
MBPP	48.40	63.40	<u>69.00</u>	69.80
CRUX-O	36.80	48.50	<u>61.10</u>	62.00
<i>Multilingual Tasks</i>				
MGSM	38.92	63.60	<u>74.68</u>	76.02
MMMLU	59.65	71.34	78.34	<u>75.72</u>
IINCLUDE	44.94	53.98	60.26	<u>59.40</u>

Претрейнинг: результаты Qwen3-4B



Table 7: Comparison among Qwen3-4B-Base and other strong open-source baselines. The highest and second-best scores are shown in bold and underlined, respectively.

	Gemma-3-4B Base	Qwen2.5-3B Base	Qwen2.5-7B Base	Qwen3-4B Base
Architecture	Dense	Dense	Dense	Dense
# Total Params	4B	3B	7B	4B
# Activated Params	4B	3B	7B	4B
<i>General Tasks</i>				
MMLU	59.51	65.62	74.16	<u>72.99</u>
MMLU-Redux	56.91	63.68	<u>71.06</u>	72.79
MMLU-Pro	29.23	34.61	<u>45.00</u>	50.58
SuperGPQA	17.68	20.31	<u>26.34</u>	28.43
BBH	51.70	56.30	<u>70.40</u>	72.59
<i>Math & STEM Tasks</i>				
GPQA	24.24	26.26	<u>36.36</u>	36.87
GSM8K	43.97	79.08	<u>85.36</u>	87.79
MATH	26.10	42.64	<u>49.80</u>	54.10
<i>Coding Tasks</i>				
EvalPlus	43.23	46.28	<u>62.18</u>	63.53
MultiPL-E	28.06	39.65	<u>50.73</u>	53.13
MBPP	46.40	54.60	<u>63.40</u>	67.00
CRUX-O	34.00	36.50	<u>48.50</u>	55.00
<i>Multilingual Tasks</i>				
MGSM	33.11	47.53	<u>63.60</u>	67.74
MMMLU	59.62	65.55	<u>71.34</u>	71.42
INCLUDE	49.06	45.90	<u>53.98</u>	56.29

Претрейнинг: итога



- Qwen3-235B-A22B составляет конкуренцию DeepSeekV3, имея в 3 раза меньше параметров
 - Хватает 1 узла с 8A100 для того, чтобы инферить в BF16!
- Qwen3-32B где-то между Qwen2.5-32B и Qwen2.5-72B
 - неплохой рост в multilingual
- Qwen3-30B-A3B +- на уровне именно с Qwen3-14B
 - Занимает 30B памяти, но очень быстрая за счет 3B активных
- Qwen3-8B на уровне Qwen2.5-14B, а Qwen3-4B, на уровне Qwen2.5-7B!

Но помним, бенчмарки не всегда отражают реальное положение вещей!

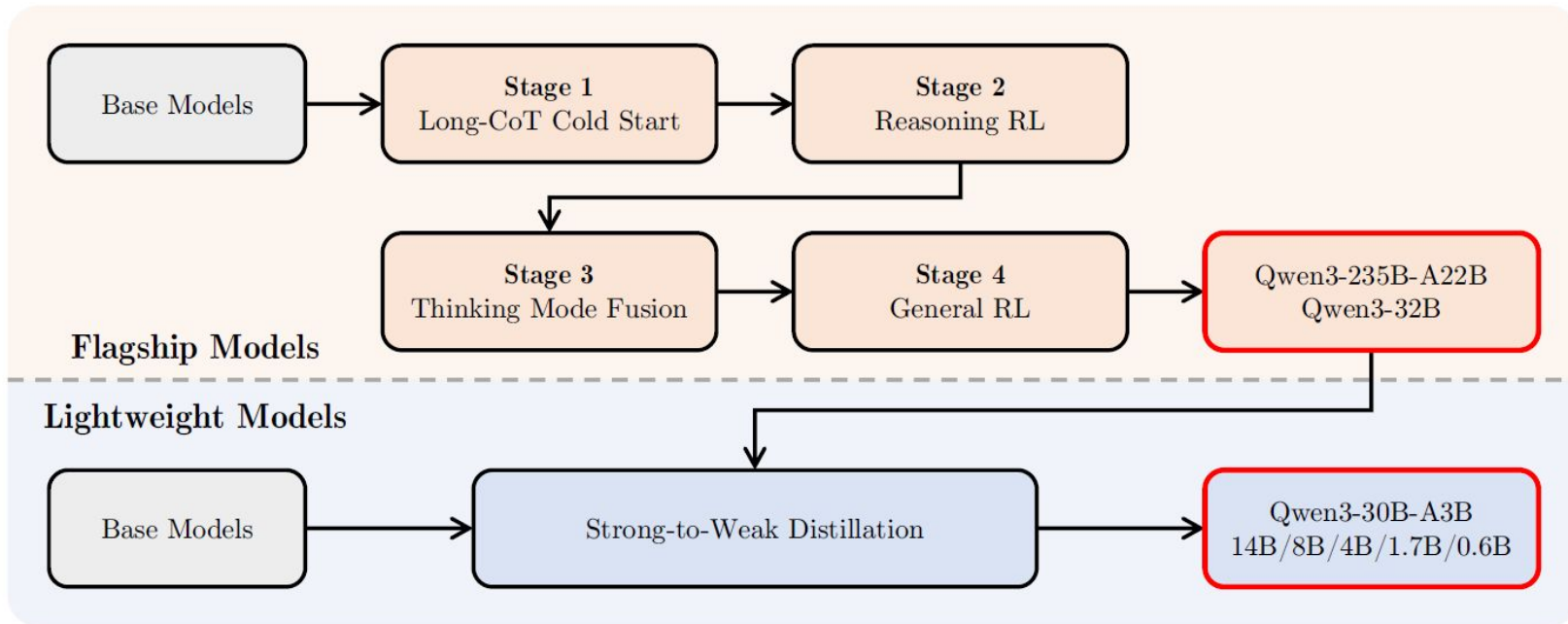


Figure 1: Post-training pipeline of the Qwen3 series models.



- Гибридный ризонинг - модели могут как размышлять, так и сразу отвечать
 - /no_think токен в конец user content части - модель не будет размышлять
 - можно в рамках одного диалога использовать оба варианта в перемешку!
- Strong-to-weak дистилляция
 - Сначала для топовых моделей (32B и 235B-A22B) провели многоступенчатую схему обучения
 - Затем дистиллировали в меньшие модели в 2 шага: дообучение на генерациях и дистилляция на логитах (kl div)



Шаги дообучения

- Stage-1: Long-CoT Cold Start
 - По сути SFT на цепочках (QWQ и Qwen2.5-72B-Instruct)
- Stage-2: Reasoning RL
 - GRPO
 - “3,995 query-verifier pairs”
- Stage-3: Thinking Mode Fusion
 - На деле учат (SFT) ризонинг модель работать не в ризонинг режиме, сохраняя при этом ризонинг режим
- Stage-4: General RL
 - Дополнительная “полировка” модели на различных задачах
 - Rule-based Reward, Model-based Reward with Reference Answer, Model-based Reward without Reference Answer

Влияние шагов дообучения на качество



Table 22: Performance of Qwen3-32B after Reasoning RL (Stage 2), Thinking Mode Fusion (Stage 3), and General RL (Stage 4). Benchmarks with * are in-house datasets.

		Stage 2	Stage 3		Stage 4	
		Reasoning RL	Thinking Mode Fusion		General RL	
	Benchmark	Thinking	Thinking	Non-Thinking	Thinking	Non-Thinking
<i>General Tasks</i>	LiveBench 2024-11-25	68.6	70.9 ^{+2.3}	57.1	74.9 ^{+4.0}	59.8 ^{+2.8}
	Arena-Hard	86.8	89.4 ^{+2.6}	88.5	93.8 ^{+4.4}	92.8 ^{+4.3}
	CounterFactQA*	50.4	61.3 ^{+10.9}	64.3	68.1 ^{+6.8}	66.4 ^{+2.1}
<i>Instruction & Format Following</i>	IFEval strict prompt	73.0	78.4 ^{+5.4}	78.4	85.0 ^{+6.6}	83.2 ^{+4.8}
	Multi-IF	61.4	64.6 ^{+3.2}	65.2	73.0 ^{+8.4}	70.7 ^{+5.5}
	LengthCtrl*	62.6	70.6 ^{+8.0}	84.9	73.5 ^{+2.9}	87.3 ^{+2.4}
	ThinkFollow*	-		88.7	98.9 ^{+10.2}	
<i>Agent</i>	BFCL v3	69.0	68.4 ^{-0.6}	61.5	70.3 ^{+1.9}	63.0 ^{+1.5}
	ToolUse*	63.3	70.4 ^{+7.1}	73.2	85.5 ^{+15.1}	86.5 ^{+13.3}
<i>Knowledge & STEM</i>	MMLU-Redux	91.4	91.0 ^{-0.4}	86.7	90.9 ^{-0.1}	85.7 ^{-1.0}
	GPQA-Diamond	68.8	69.0 ^{+0.2}	50.4	68.4 ^{-0.6}	54.6 ^{+4.3}
<i>Math & Coding</i>	AIME'24	83.8	81.9 ^{-1.9}	28.5	81.4 ^{-0.5}	31.0 ^{+2.5}
	LiveCodeBench v5	68.4	67.2 ^{-1.2}	31.1	65.7 ^{-1.5}	31.3 ^{+0.2}

Шаблон для thinking и non-thinking



Thinking Mode	Non-Thinking Mode
<code>< im_start >user</code> <code>{query} /think< im_end ></code> <code>< im_start >assistant</code> <code><think></code> <code>{thinking_content}</code> <code></think></code> <code>{response}< im_end ></code>	<code>< im_start >user</code> <code>{query} /no_think< im_end ></code> <code>< im_start >assistant</code> <code><think></code> <code></think></code> <code>{response}< im_end ></code>

Результаты для инструктивной Qwen3-235B-A22B



Московский
государственный
университет
имени М.В. Ломоносова

Thinking

	OpenAI-o1	DeepSeek-R1	Grok-3-Beta (Think)	Gemini2.5-Pro	Qwen3-235B-A22B
	Architecture	MoE	-	-	MoE
	# Activated Params	37B	-	-	22B
	# Total Params	671B	-	-	235B
<i>General Tasks</i>	MMLU-Redux	92.8	92.9	-	93.7
	GPQA-Diamond	78.0	71.5	80.2	84.0
	C-Eval	85.5	91.8	-	82.9
	LiveBench 2024-11-25	75.7	71.6	-	82.4
<i>Alignment Tasks</i>	IFEval strict prompt	92.6	83.3	-	89.5
	Arena-Hard	92.1	92.3	-	96.4
	AlignBench v1.1	8.86	8.76	-	9.03
	Creative Writing v3	81.7	85.5	-	86.0
	WritingBench	7.69	7.71	-	8.09
<i>Math & Text Reasoning</i>	MATH-500	96.4	97.3	-	98.8
	AIME'24	74.3	79.8	83.9	92.0
	AIME'25	79.2	70.0	77.3	86.7
	ZebraLogic	81.0	78.7	-	87.4
	AutoLogi	79.8	86.1	-	85.4
<i>Agent & Coding</i>	BFCL v3	67.8	56.9	-	62.9
	LiveCodeBench v5	63.9	64.3	70.6	70.4
	CodeForces (Rating / Percentile)	1891 / 96.7%	2029 / 98.1%	-	2001 / 97.9%
<i>Multilingual Tasks</i>	Multi-IF	48.8	67.7	-	77.8
	INCLUDE	84.6	82.7	-	85.1
	MMMLU 14 languages	88.4	86.4	-	86.9
	MT-AIME2024	67.4	73.5	-	76.9
	PolyMath	38.9	47.1	-	52.2
	MLogiQA	75.5	73.8	-	75.6

Non-Thinking

	GPT-4o -2024-11-20	DeepSeek-V3	Qwen2.5-72B -Instruct	LLaMA-4 -Maverick	Qwen3-235B-A22B
	Architecture	MoE	Dense	MoE	MoE
	# Activated Params	37B	72B	17B	22B
	# Total Params	671B	671B	402B	235B
<i>General Tasks</i>	MMLU-Redux	87.0	89.1	86.8	91.8
	GPQA-Diamond	46.0	59.1	49.0	69.8
	C-Eval	75.5	86.5	84.7	83.5
	LiveBench 2024-11-25	52.2	60.5	51.4	59.5
<i>Alignment Tasks</i>	IFEval strict prompt	86.5	86.1	84.1	86.7
	Arena-Hard	85.3	85.5	81.2	82.7
	AlignBench v1.1	8.42	8.64	7.89	7.97
	Creative Writing v3	81.1	74.0	61.8	61.3
	WritingBench	7.11	6.49	7.06	5.46
<i>Math & Text Reasoning</i>	MATH-500	77.2	90.2	83.6	90.6
	AIME'24	11.1	39.2	18.9	38.5
	AIME'25	7.6	28.8	15.0	15.9
	ZebraLogic	27.4	42.1	26.6	40.0
	AutoLogi	65.9	76.1	66.1	75.2
<i>Agent & Coding</i>	BFCL v3	72.5	57.6	63.4	52.9
	LiveCodeBench v5	32.7	33.1	30.7	37.2
	CodeForces (Rating / Percentile)	864 / 35.4%	1134 / 54.1%	859 / 35.0%	712 / 24.3%
<i>Multilingual Tasks</i>	Multi-IF	65.6	55.6	65.3	75.5
	INCLUDE	78.8	76.7	69.6	80.9
	MMMLU 14 languages	80.3	81.1	76.9	82.5
	MT-AIME2024	9.2	20.9	12.7	27.0
	PolyMath	13.7	20.4	16.9	26.1
	MLogiQA	57.4	58.9	59.3	59.9

Результаты для инструктивной Qwen3-32B



Thinking

	DeepSeek-R1 -Distill-Llama-70B	QwQ-32B	OpenAI-o3-mini (medium)	Qwen3-32B
	Dense	Dense	-	Dense
	70B	32B	-	32B
	70B	32B	-	32B
<i>General Tasks</i>				
	89.3	<u>90.0</u>	<u>90.0</u>	<u>90.9</u>
	65.2	<u>65.6</u>	<u>76.8</u>	<u>68.4</u>
	71.8	<u>88.4</u>	<u>75.1</u>	<u>87.3</u>
	54.5	<u>72.0</u>	<u>70.0</u>	<u>74.9</u>
<i>Alignment Tasks</i>				
	79.3	<u>83.9</u>	<u>91.5</u>	<u>85.0</u>
	60.6	<u>89.5</u>	<u>89.0</u>	<u>93.8</u>
	6.74	<u>8.70</u>	<u>8.38</u>	<u>8.72</u>
	62.1	<u>82.4</u>	<u>74.8</u>	<u>81.0</u>
	6.08	<u>7.86</u>	<u>7.52</u>	<u>7.90</u>
<i>Math & Text Reasoning</i>				
	94.5	<u>98.0</u>	<u>98.0</u>	<u>97.2</u>
	70.0	<u>79.5</u>	<u>79.6</u>	<u>81.4</u>
	56.3	<u>69.5</u>	<u>74.8</u>	<u>72.9</u>
	71.3	<u>76.8</u>	<u>88.9</u>	<u>88.8</u>
	83.5	<u>88.1</u>	<u>86.3</u>	<u>87.3</u>
<i>Agent & Coding</i>				
	49.3	<u>66.4</u>	<u>64.6</u>	<u>70.3</u>
	54.5	<u>62.7</u>	<u>66.3</u>	<u>65.7</u>
	1633 / 91.4%	<u>1982 / 97.7%</u>	<u>2036 / 98.1%</u>	<u>1977 / 97.7%</u>
<i>Multilingual Tasks</i>				
	57.6	<u>68.3</u>	<u>48.4</u>	<u>73.0</u>
	62.1	<u>69.7</u>	<u>73.1</u>	<u>73.7</u>
	69.6	<u>80.9</u>	<u>79.3</u>	<u>80.6</u>
	29.3	<u>68.0</u>	<u>73.9</u>	<u>75.0</u>
	29.4	<u>45.9</u>	<u>38.6</u>	<u>47.4</u>
	60.3	<u>75.5</u>	<u>71.1</u>	<u>76.3</u>

Non-Thinking

	GPT-4o-mini -2024-07-18	LLaMA-4 -Scout	Qwen2.5-72B -Instruct	Qwen3-32B
	-	MoE	Dense	Dense
	-	17B	72B	32B
	-	109B	72B	32B
<i>General Tasks</i>				
	81.5	<u>86.3</u>	<u>86.8</u>	85.7
	40.2	<u>57.2</u>	<u>49.0</u>	54.6
	66.3	<u>78.2</u>	<u>84.7</u>	<u>83.3</u>
	41.3	<u>47.6</u>	<u>51.4</u>	<u>59.8</u>
<i>Alignment Tasks</i>				
	80.4	<u>84.7</u>	<u>84.1</u>	83.2
	74.9	<u>70.5</u>	<u>81.2</u>	<u>92.8</u>
	7.81	<u>7.49</u>	<u>7.89</u>	<u>8.58</u>
	70.3	<u>55.0</u>	<u>61.8</u>	<u>78.3</u>
	5.98	<u>5.49</u>	<u>7.06</u>	<u>7.54</u>
<i>Math & Text Reasoning</i>				
	78.2	<u>82.6</u>	<u>83.6</u>	88.6
	8.1	<u>28.6</u>	<u>18.9</u>	31.0
	8.8	<u>10.0</u>	<u>15.0</u>	20.2
	20.1	<u>24.2</u>	<u>26.6</u>	29.2
	52.6	<u>56.8</u>	<u>66.1</u>	78.5
<i>Agent & Coding</i>				
	64.0	<u>45.4</u>	<u>63.4</u>	63.0
	27.9	<u>29.8</u>	<u>30.7</u>	31.3
	1113 / 52.6%	<u>981 / 43.7%</u>	<u>859 / 35.0%</u>	<u>1353 / 71.0%</u>
<i>Multilingual Tasks</i>				
	62.4	<u>64.2</u>	<u>65.3</u>	70.7
	66.0	<u>74.1</u>	<u>69.6</u>	70.9
	72.1	<u>77.5</u>	<u>76.9</u>	76.5
	6.0	<u>19.1</u>	<u>12.7</u>	24.1
	12.0	<u>20.9</u>	<u>16.9</u>	22.5
	42.6	<u>53.9</u>	<u>59.3</u>	62.9

Влияние количества размышлений на итоговое качество

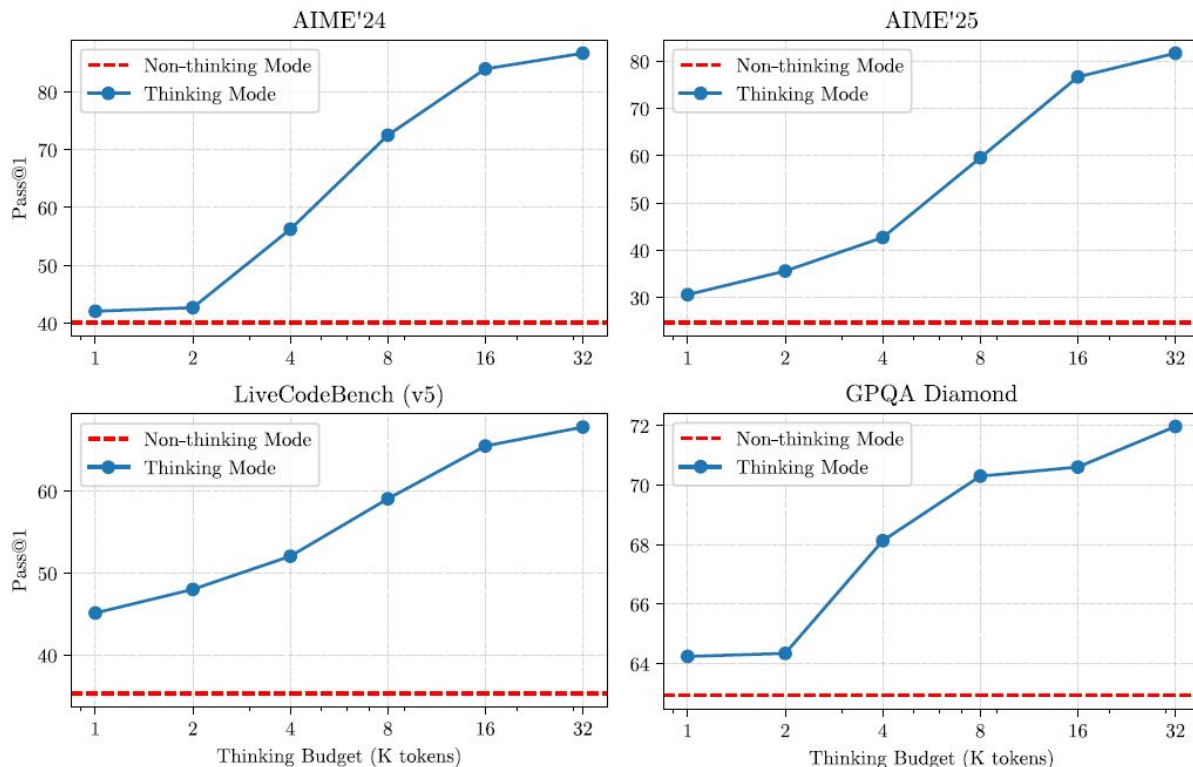


Figure 2: Performance of Owen3-235B-A22B with respect to the thinking budget.

Дистилляция > RL



Table 21: Comparison of reinforcement learning and on-policy distillation on Qwen3-8B. Numbers in parentheses indicate pass@64 scores.

Method	AIME'24	AIME'25	MATH500	LiveCodeBench v5	MMLU -Redux	GPQA -Diamond	GPU Hours
Off-policy Distillation	55.0 (90.0)	42.8 (83.3)	92.4	42.0	86.4	55.6	-
+ Reinforcement Learning	67.6 (90.0)	55.5 (83.3)	94.8	52.9	86.9	61.3	17,920
+ On-policy Distillation	74.4 (93.3)	65.5 (86.7)	97.0	60.3	88.3	63.3	1,800

Off-policy - модель дообучалась на генерациях

On-policy - модель дообучалась дополнительно на логитах БОльшей модели!

Метрики на русском языке



Table 32: **Benchmark scores for language: Russian (ru)**. The highest and second-best scores are shown in **bold** and underlined, respectively.

	Model	Multi-IF	INCLUDE	MT-AIME24	PolyMath	Average
<i>Thinking Mode</i>	Gemini2.5-Pro	68.1	80.4	70.0	<u>52.3</u>	<u>67.7</u>
	QwQ-32B	61.2	73.2	<u>76.7</u>	43.6	63.7
	Qwen3-235B-A22B	62.2	80.4	80.0	53.1	68.9
	Qwen3-32B	62.5	73.2	63.3	46.5	61.4
	Qwen3-30B-A3B	60.7	<u>76.8</u>	73.3	45.4	64.0
	Qwen3-14B	<u>63.6</u>	80.4	66.7	46.4	64.3
	Qwen3-8B	<u>62.9</u>	69.6	63.3	37.7	58.4
	Qwen3-4B	52.8	69.6	56.7	36.6	53.9
	Qwen3-1.7B	37.8	46.4	20.0	22.8	31.8
	Qwen3-0.6B	26.4	46.4	3.3	7.0	20.8
<i>Non-thinking Mode</i>	GPT-4o-2024-1120	52.0	80.4	20.0	13.7	41.5
	Gemma-3-27b-IT	57.3	71.4	23.3	21.6	43.4
	Qwen2.5-72B-Instruct	54.1	67.9	20.0	13.3	38.8
	Qwen3-235B-A22B	56.7	<u>75.0</u>	40.0	26.1	49.4
	Qwen3-32B	58.6	71.4	<u>30.0</u>	23.3	<u>45.8</u>
	Qwen3-30B-A3B	58.0	73.2	<u>30.0</u>	21.1	45.6
	Qwen3-14B	60.3	71.4	26.7	<u>24.2</u>	45.6
	Qwen3-8B	<u>59.3</u>	58.9	20.0	22.8	40.2
	Qwen3-4B	46.1	58.9	13.3	17.8	34.0
	Qwen3-1.7B	34.8	41.1	3.3	13.2	23.1
Qwen3-0.6B	25.5	46.4	0.0	5.8	19.4	



- Модели вышли не так давно и вопрос, насколько качество реально хорошее
 - Но первые впечатления отличные, хотя шаг меньше от 2.5, чем хотелось бы
- Модели очень “плотные” на знания, дообучать сложно
 - Попытки дообучить на русских инструктивных данных приводят к деградации (пока что)
 - Одна из причин - дистилляция была одним из последних шагов, обычный SFT “затирает” навыки?
 - Вторая причина - средняя длина ответа в `no_think` - 700+ токенов! У версии 2.5 было ~420!

Discussion v2



- Отсутствуют датасеты инструкций с гибридным режимом
 - Но, его можно сделать из датасетов с ризонингом
 - Уже начали генерировать синтетику через Qwen3-235B-A22B
- Не “релизались” на lmsys заранее, как это делают в последнее время!
 - Случай с LLaMa-4 очень яркий негативный пример переучивания на арену

Imsys Qwen3-235B-A22B



Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License
5	3	o4-mini-2025-04-16	1324	+9/-5	5083	OpenAI	Proprietary
5	2	DeepSeek-V3-0324	1323	+7/-5	8753	DeepSeek	MIT
8	2	o1-2024-12-17	1319	+3/-4	29036	OpenAI	Proprietary
8	0	DeepSeek-R1	1316	+5/-3	18493	DeepSeek	MIT
10	14	Claude-3.7-Sonnet...(thinking-32k)	1312	+5/-6	12038	Anthropic	Proprietary
13	1	o1-preview	1303	+3/-3	33171	OpenAI	Proprietary
13	18	Claude-3.7-Sonnet	1298	+4/-4	17387	Anthropic	Proprietary
13	-5	Hunyuan-Turbo-20250416	1295	+9/-10	3699	Tencent	Proprietary
14	4	GPT-4.1-mini-2025-04-14	1293	+7/-11	4950	OpenAI	Proprietary
14	-4	Qwen3-235B-A22B	1286	+10/-8	3611	Alibaba	Apache 2.0
15	3	o3-mini-high	1287	+4/-4	19410	OpenAI	Proprietary
15	-3	Qwen2.5-Max	1286	+4/-3	23180	Alibaba	Proprietary
15	22	Claude-3.5-Sonnet...(20241022)	1286	+2/-3	65435	Anthropic	Proprietary
15	-7	Gemini-2.0-Flash-001	1285	+3/-4	24913	Google	Proprietary
15	-3	Gemma-3-27B-it	1281	+4/-5	12343	Google	Gemma
21	-3	DeepSeek-V3	1277	+3/-4	22838	DeepSeek	DeepSeek
23	-4	Gemini-2.0-Flash-Lite	1270	+4/-3	25020	Google	Proprietary
23	-2	o3-mini	1269	+5/-4	24918	OpenAI	Proprietary