南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Advanced Natural Language Processing

Lecture 12: LLM Data Synthesis

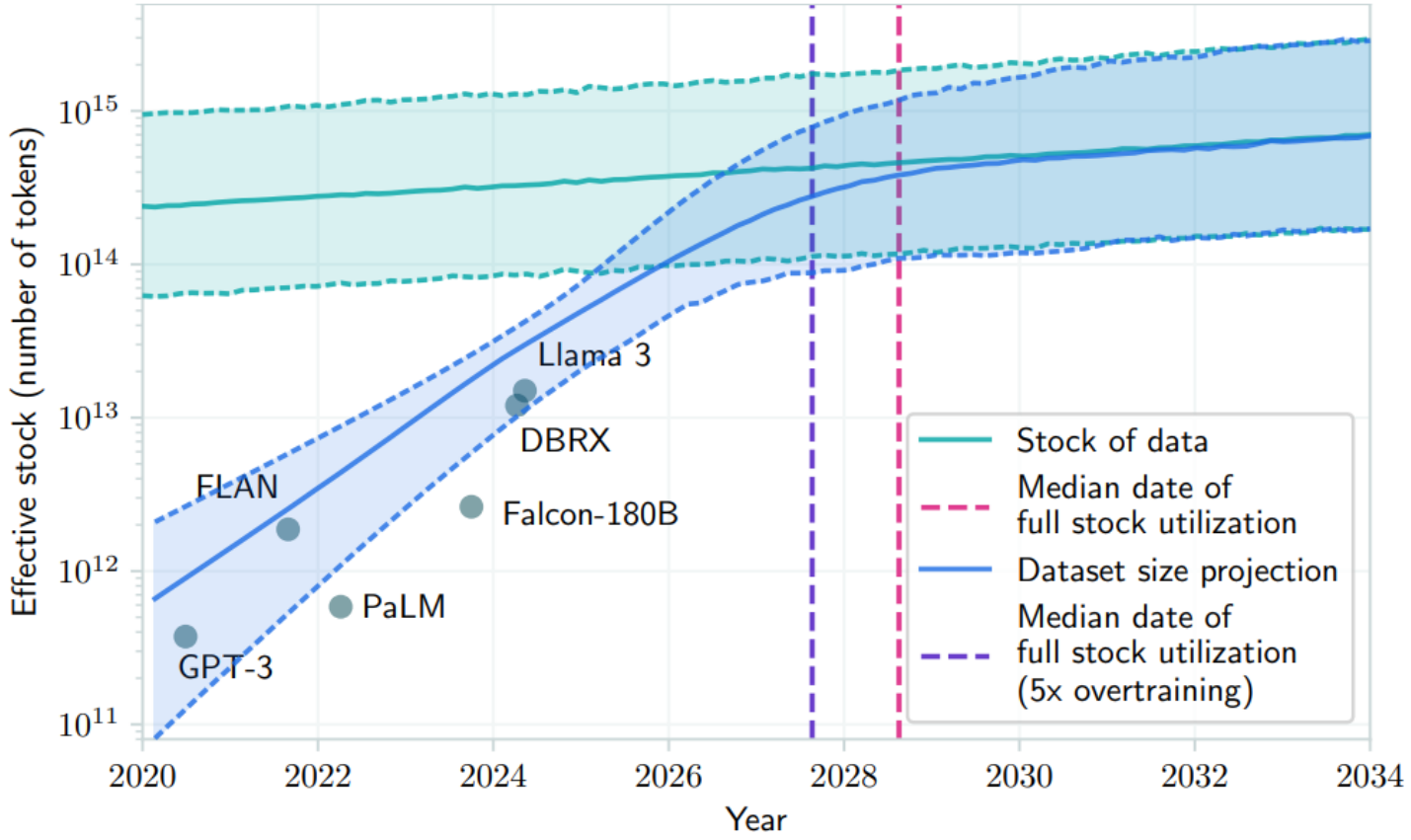
陈冠华 CHEN Guanhua

Department of Statistics and Data Science

Background

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- In the next 10 years, the rate of data growth will not be able to support the expansion of LLMs



Will we run out of data? Limits of LLM scaling based on human-generated data (ICML 2024)

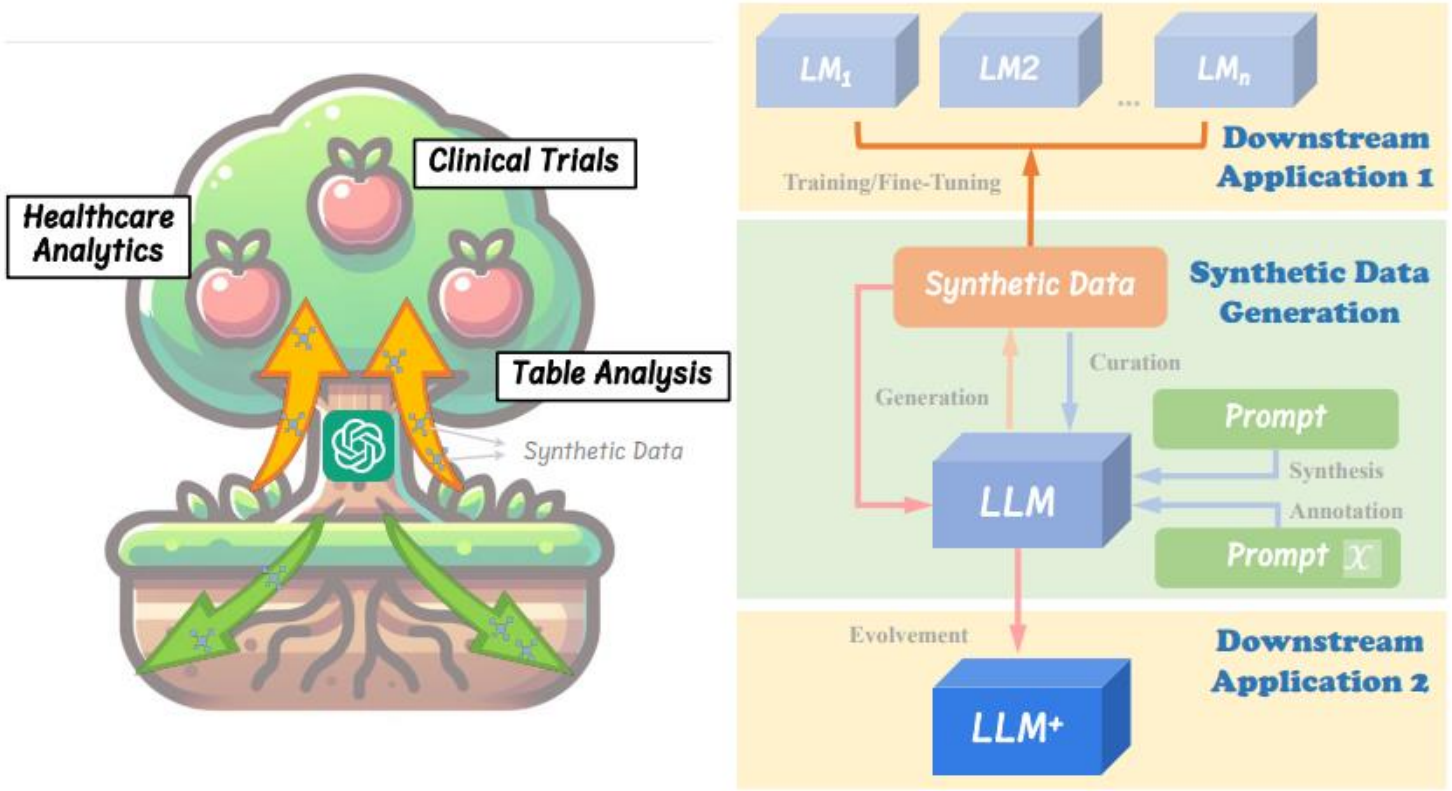
Background

Human-labeled data

- Advantages
  - High quality and reliability
  - Better alignment with human preference
- Disadvantages
  - Expensive
  - Time-consuming / low scalability

Background

- Synthetic data
  - Aim to be effective and relatively low-cost alternative of real data



Best Practices and Lessons Learned on Synthetic Data (COLM 2024)

On LLMs-Driven Synthetic Data Generation, Curation, and Evaluation: A Survey

Background

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Where to use synthetic data
  - Pretraining, continuous pretraining
  - SFT/instruction tuning
  - RLHF/alignment tuning
  - Agentic applications
  - Multimodal/embodied tasks

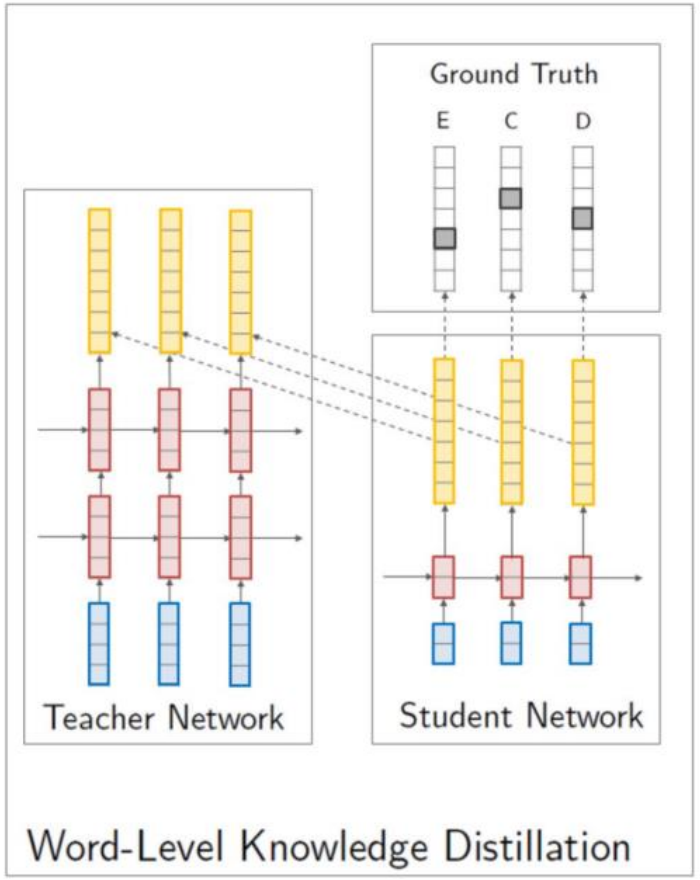
Approaches to Synthesize Data

- Sampling-based generation
  - Generating instruction data from scratch
  - Generating instruction data from document
- Transformation of existing data
- Human-AI collaboration

Background: Knowledge Distillation

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Train student model to mimic the teacher's predicted probability distribution (e.g., over words)



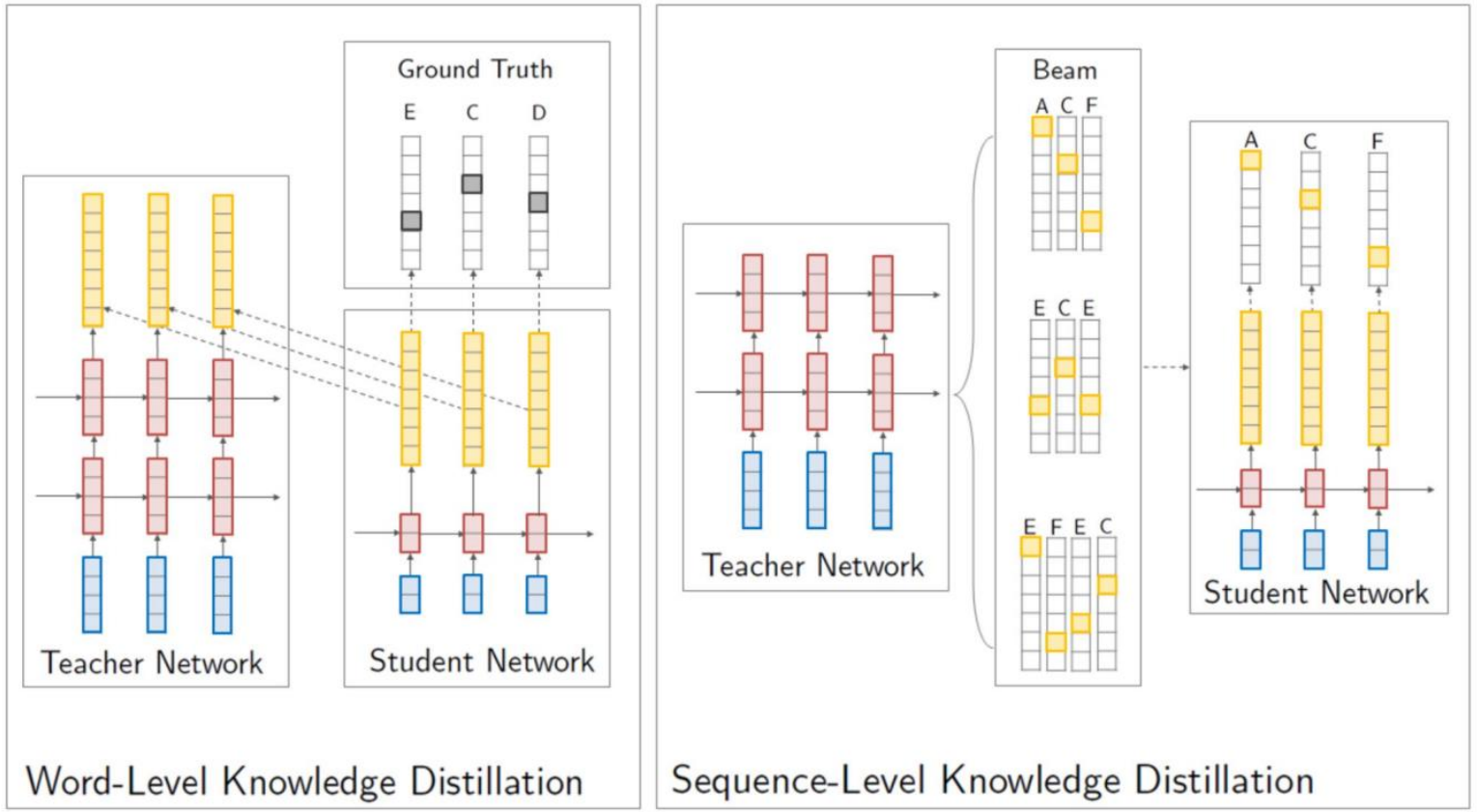
"Once the cumbersome model has been trained, we can then use a different kind of training, which we call "distillation" to transfer the knowledge to a small model"

https://synth-data-acl.github.io/static/slides/slides.pdf

Background: Knowledge Distillation

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Train student on complete generations (i.e., sequences of words) from the teacher



https://synth-data-acl.github.io/static/slides/slides.pdf

Sampling-Based Data Generation

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Generate data from an LLM for training another LM
  - Use GPT-4's in-context learning ability to generate new examples of arbitrary tasks

```
Task: Write two sentences that mean
the same thing

Sentence 1: A man is playing the flute
Sentence 2: He's playing the flute
```

Create sentence-similarity examples by prompting the model to write similar (or dissimilar) sentences

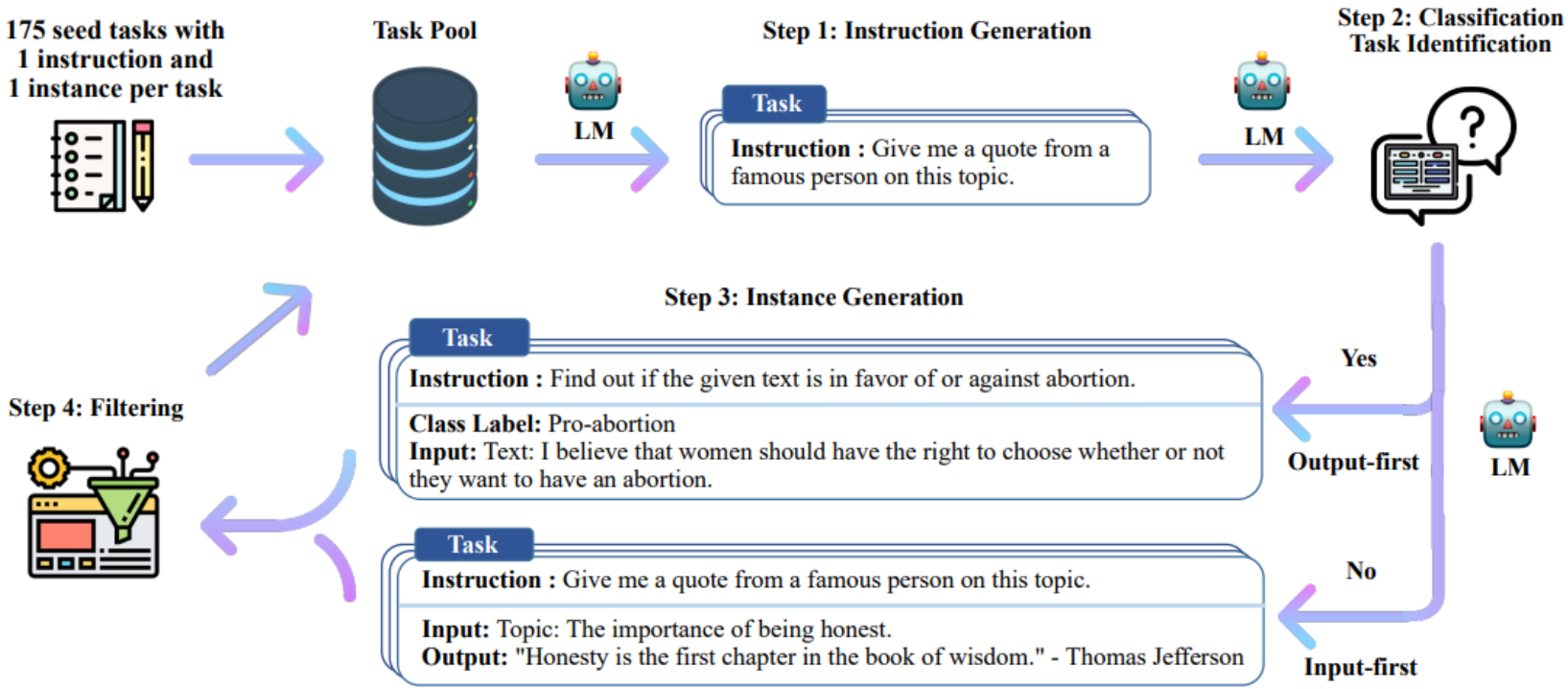
Generating Datasets with Pretrained Language Models

Generating Instruction Data From Scratch

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Sampling-based data generation can get a response, but what if we need to expand the QA pairs and need more questions?

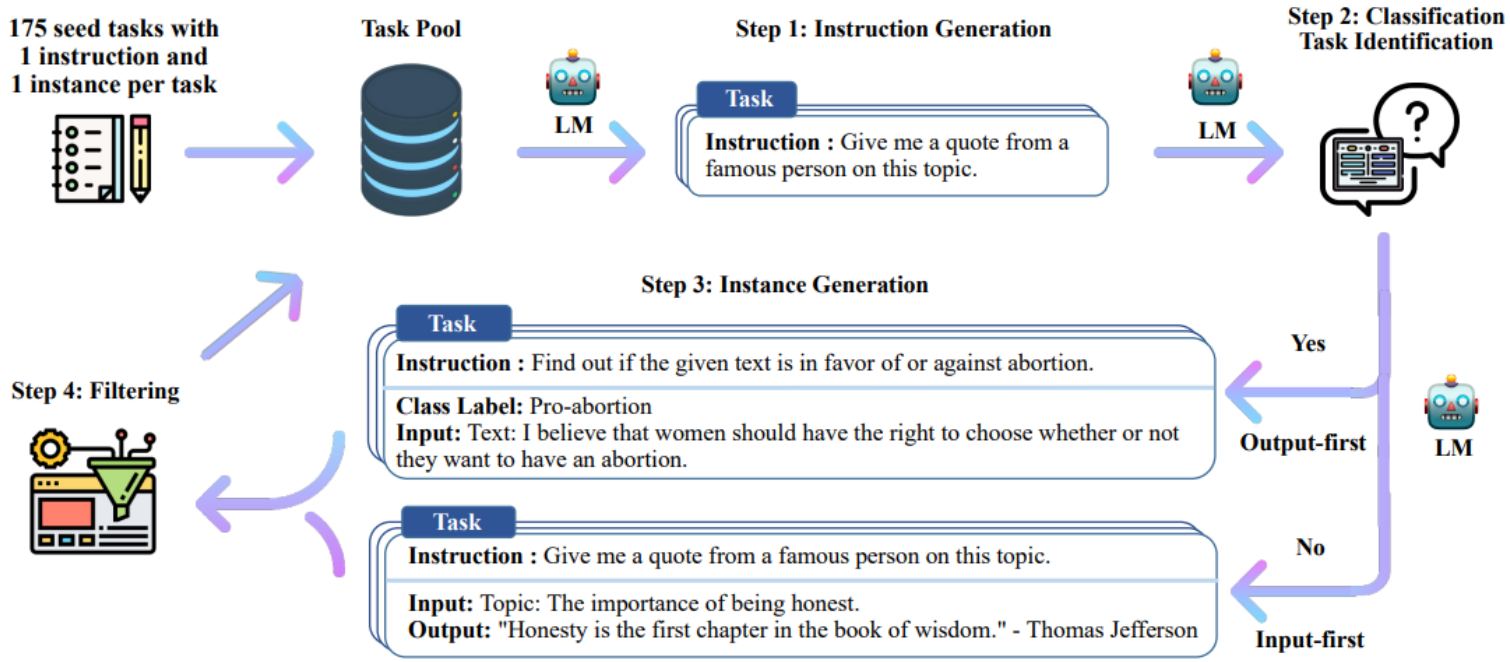
From just 175 seed examples → ~100K new examples



Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Self-Instruct

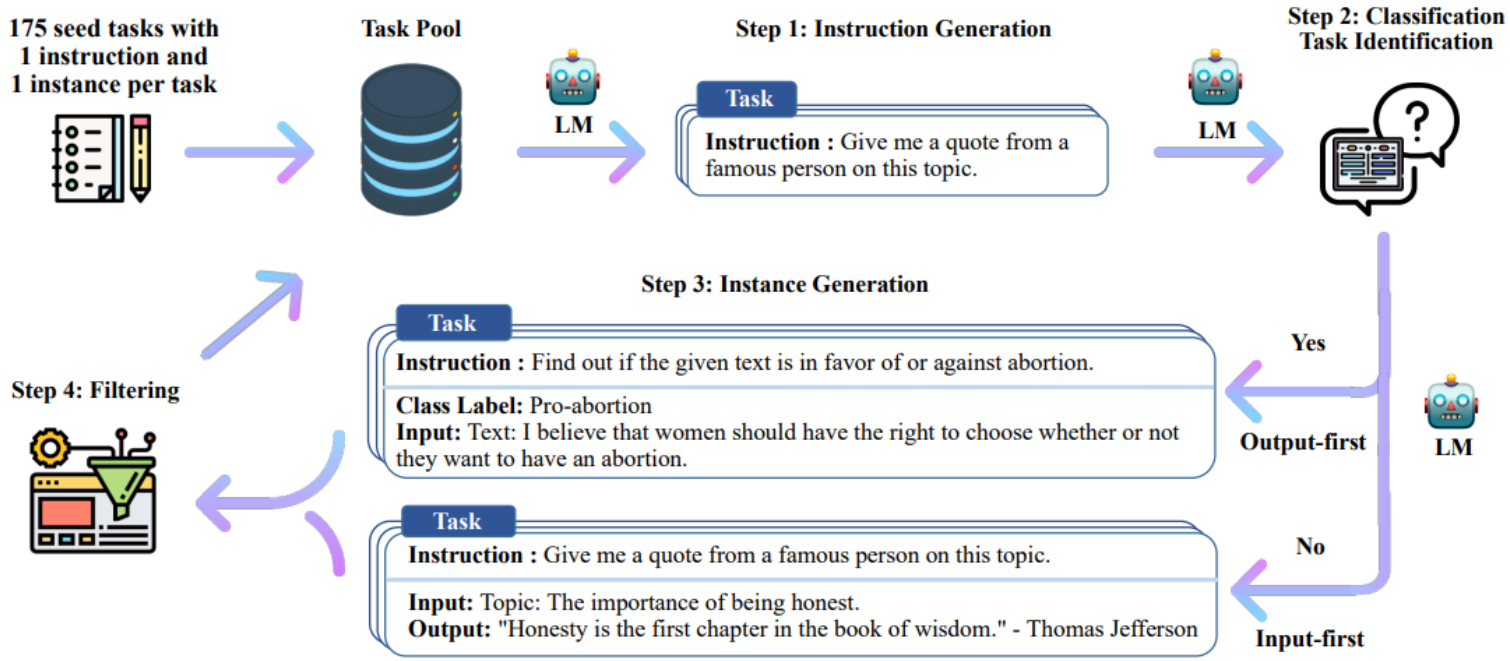
- Step 1: Instruction Generation
- Step 2: Classification Task Identification
- Step 3: Instance Generation
- Step 4: Filtering and Quality Control



Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Self-Instruct

- Step 1: Instruction Generation
  - Begins with a small seed set of 175 manually written instructions
  - In each iteration, the language model generates new instructions using in-context learning

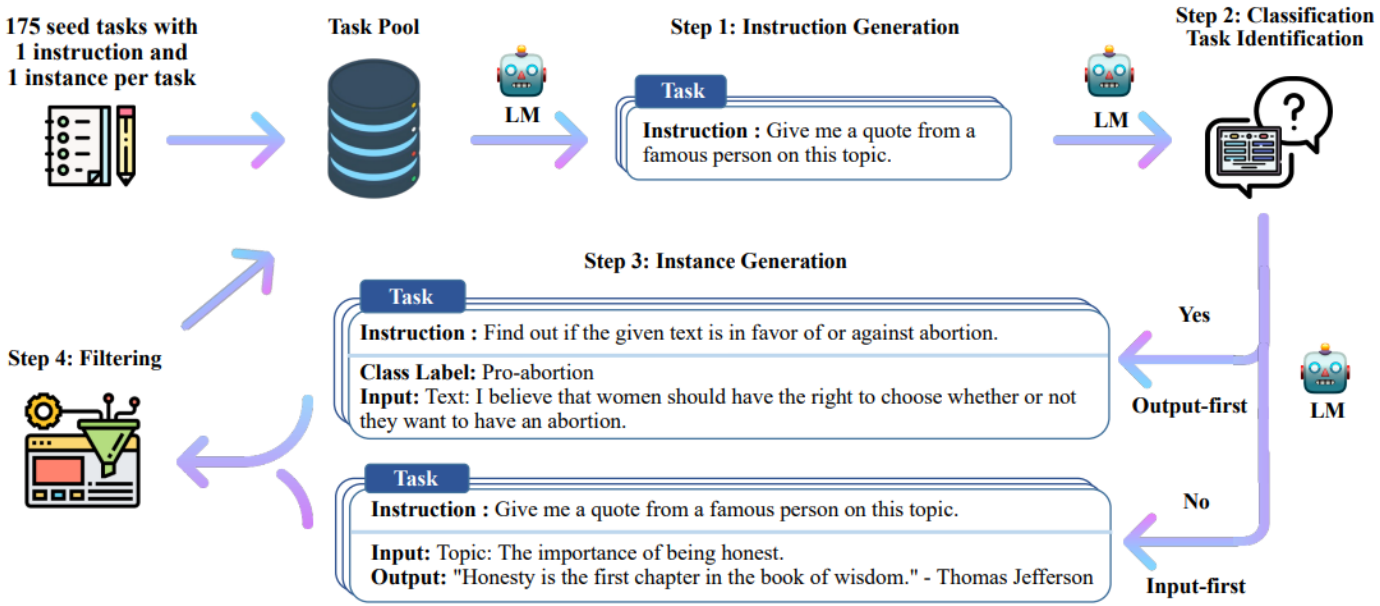


Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Self-Instruct

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

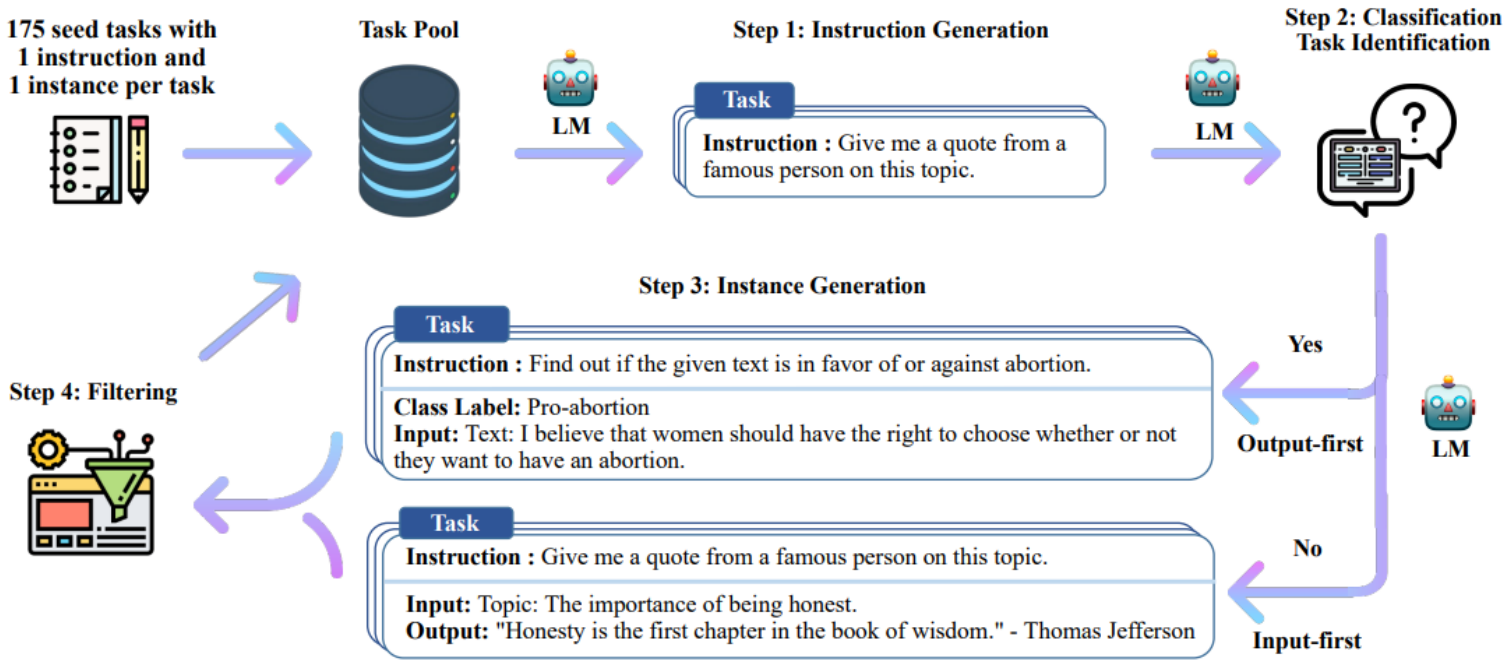
- Step 2: Classification Task Identification
  - Newly generated instructions are classified to determine whether they represent classification tasks.
  - The model uses few-shot learning to make this determination, comparing new instructions against examples of classification and non-classification tasks.



Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Self-Instruct

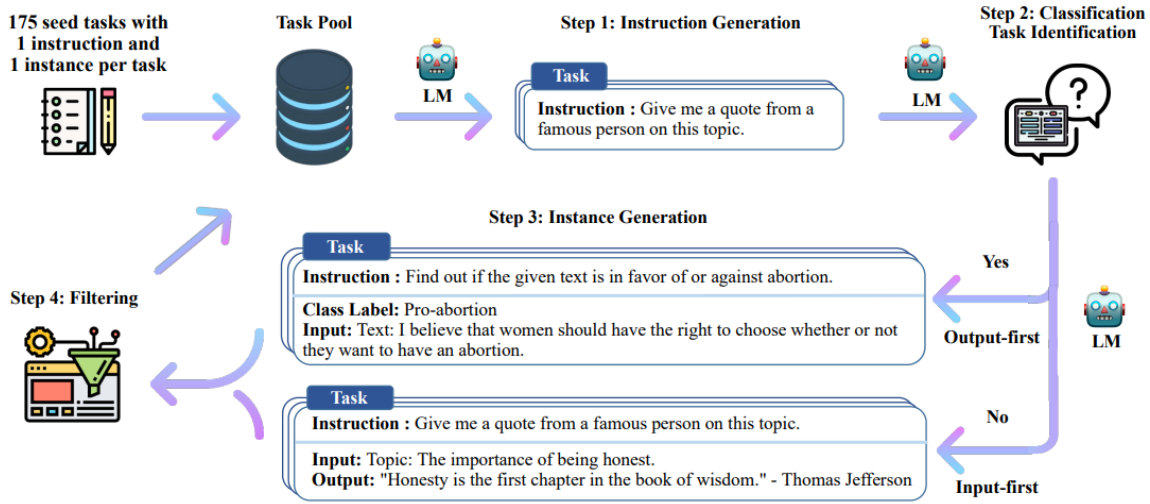
- Step 3: Instance Generation
  - Non-classification tasks
    - Input-first approach
  - Classification tasks
    - Output-first approach



Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Self-Instruct

- Step 4: Filtering and Quality Control
  - Instruction filtering: New instructions are only added if their similarity to existing instructions (measured by ROUGE-L) is below 0.7
  - Instance filtering: Exact duplicates and instances with identical inputs but different outputs are removed
  - Heuristic filtering: Invalid generations (such as overly long or repetitive outputs) are identified and discarded



Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

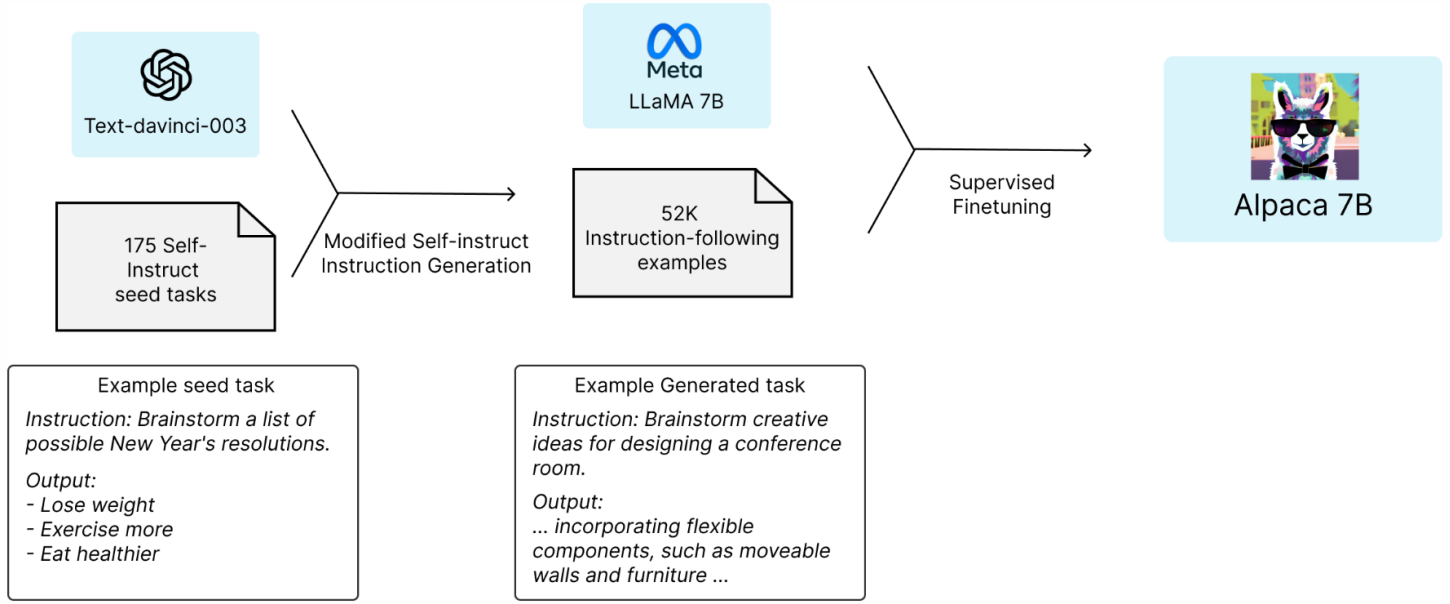
Self-Instruct

- Advantages
  - Cost-effective data generation
  - Strong empirical results
  - Transparency and reproducibility
- Disadvantages
  - Data quality issues
  - Label imbalance for classification tasks
  - Dependence on large models

Self-Instruct: Aligning Language Models with Self-Generated Instructions (ACL 2023)

Alpaca

- Alpaca: A Strong, Replicable Instruction-Following Model
- Methods：Self-Instruct
- Teacher Model：text-davinci-003
- Base Model：Llama 7B
- Data Size：52K unique instructions and the corresponding outputs



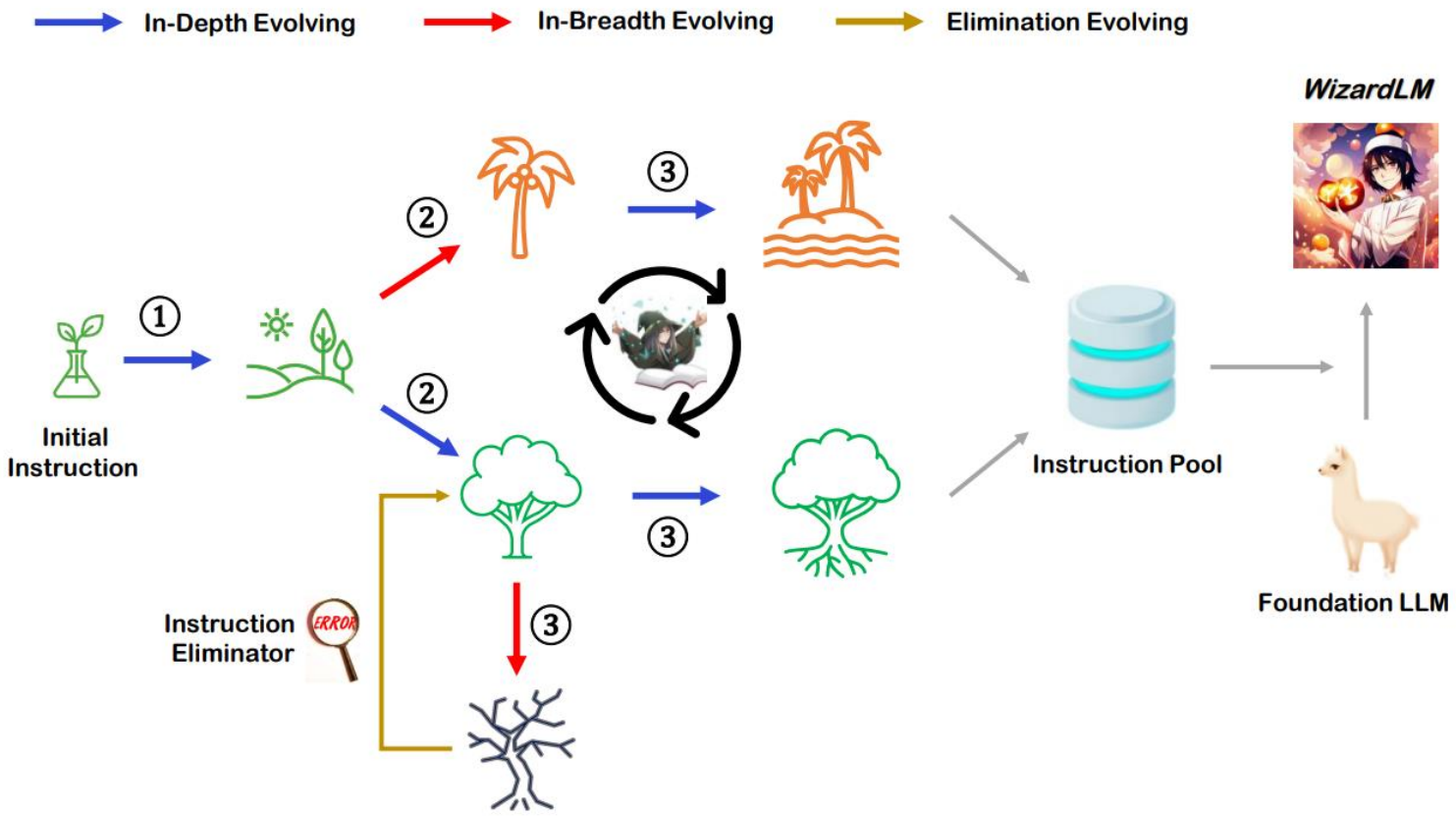
Takeaways from early efforts

- Synthetic data can reflect creativity & diversity difficult to elicit from crowd-workers
- Diversity can be more valuable than correctness!
- Synthetic data can sometimes enable self-improvement
- Data creation becomes a complex pipeline

https://synth-data-acl.github.io/static/slides/slides.pdf

WizardLM

- The problems generated by self-instruct are relatively monotonous. How to expand the problem?



WizardLM: Empowering large pre-trained language models to follow complex instructions

WizardLM

- Step 1: Initialize with Seed Instructions
- Step 2: Instruction Evolution Process
  - **In-Depth Evolving**: Make instructions more complex and difficult using 5 operations:
    - Add constraints
    - Deepening
    - Concretizing (replacing general concepts with specific ones)
    - Increase reasoning steps
    - Complicate input (adding structured data like XML, JSON, et al.)

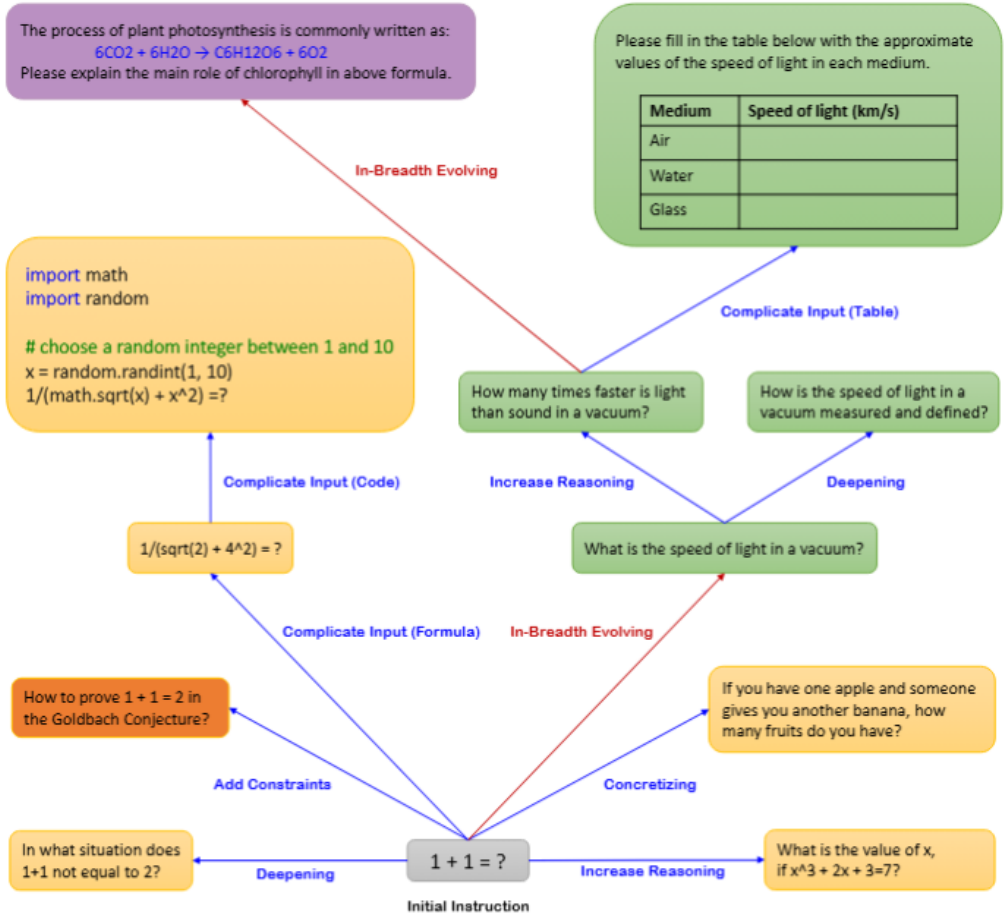


Figure 1: Running Examples of *Evol-Instruct*.

WizardLM: Empowering large pre-trained language models to follow complex instructions

WizardLM

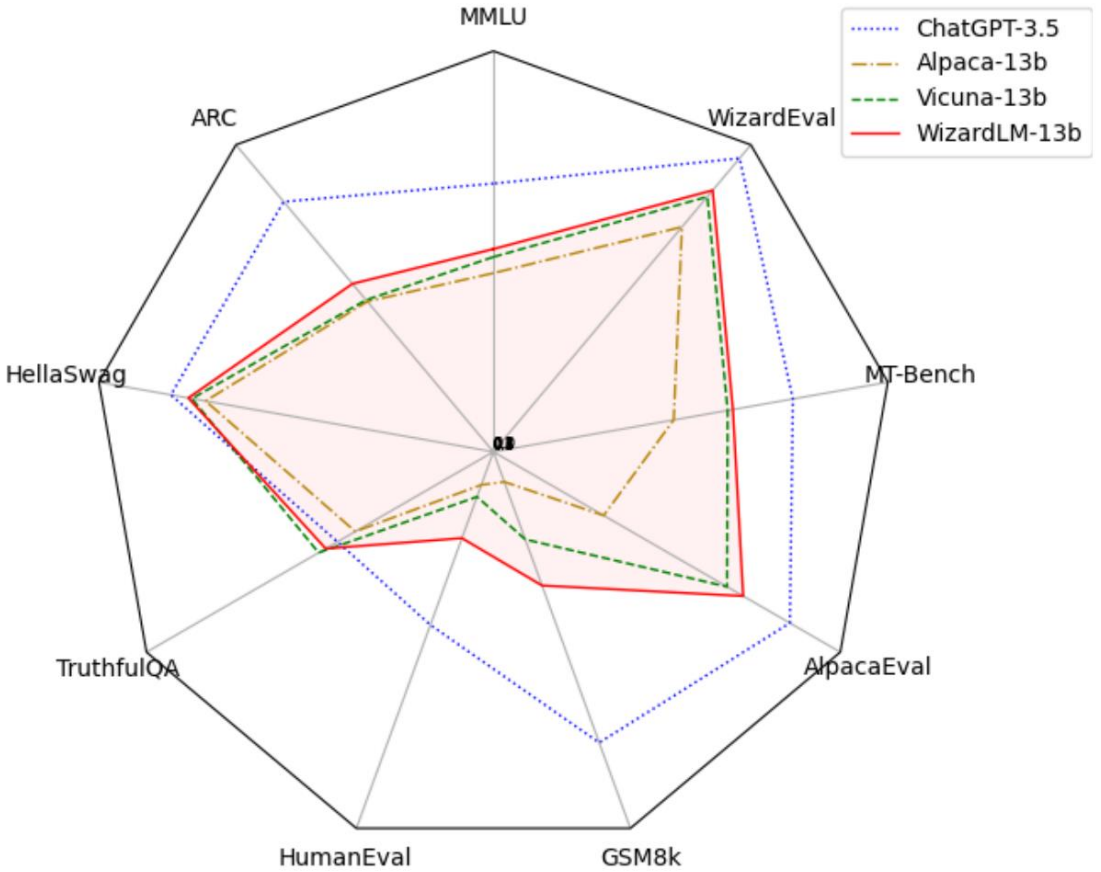
南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Step 2: Instruction Evolution Process
  - In-Breadth Evolving: Generate completely new instructions in the same domain to enhance diversity and topic coverage
  - Elimination Evolving: Filter out failed evolutions using 4 criteria:
    - No information gain vs. original
    - LLM struggles to generate response
    - Response contains only punctuation/stop words
    - Instruction copies prompt template words
- Step 3: Response Generation & Dataset Construction

WizardLM: Empowering large pre-trained language models to follow complex instructions

WizardLM

- Comprehensive abilities
  - Math, coding
  - Instruction Following skills



WizardLM: Empowering large pre-trained language models to follow complex instructions

WizardLM

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

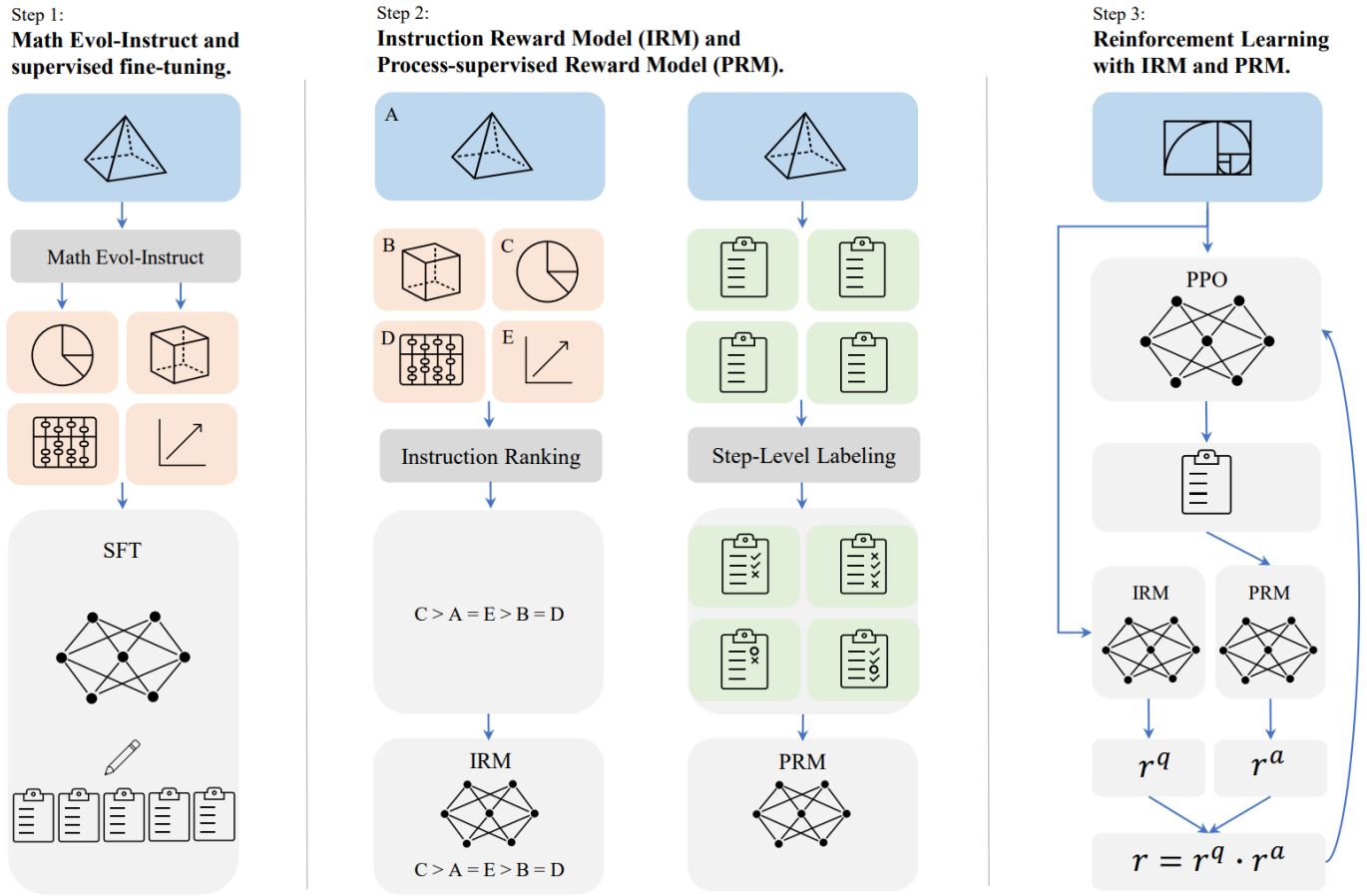
- Advantages
  - An automated approach using LLMs to evolve simple instructions into complex ones
  - Strong Performance
- Disadvantages
  - Evaluation constraints: Limited scalability and reliability of GPT-4 and human evaluation methods
  - Complexity validation: Difficulty scoring relies on LLM judgment rather than objective metrics

WizardLM: Empowering large pre-trained language models to follow complex instructions

WizardMath

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- WizardMath: Empowering Mathematical Reasoning for Large Language Models via Reinforced Evol-Instruct (ICLR 2025 Oral)



Application of WizardLM in mathematics

WizardMath

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Step 1: Math Evol-Instruct and Supervised Fine-Tuning
  - Apply upward and downward evolution to GSM8k and MATH datasets
    - Upward: Increase complexity
    - Downward: Simplify problems for diversity
  - Generate evolved instructions and use GPT-4 to create step-by-step solutions
  - Train the model via Supervised Fine-Tuning (SFT)

WizardMath: Empowering Mathematical Reasoning for
Large Language Models via Reinforced Evol-Instruct (ICLR 2025 Oral)

WizardMath

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

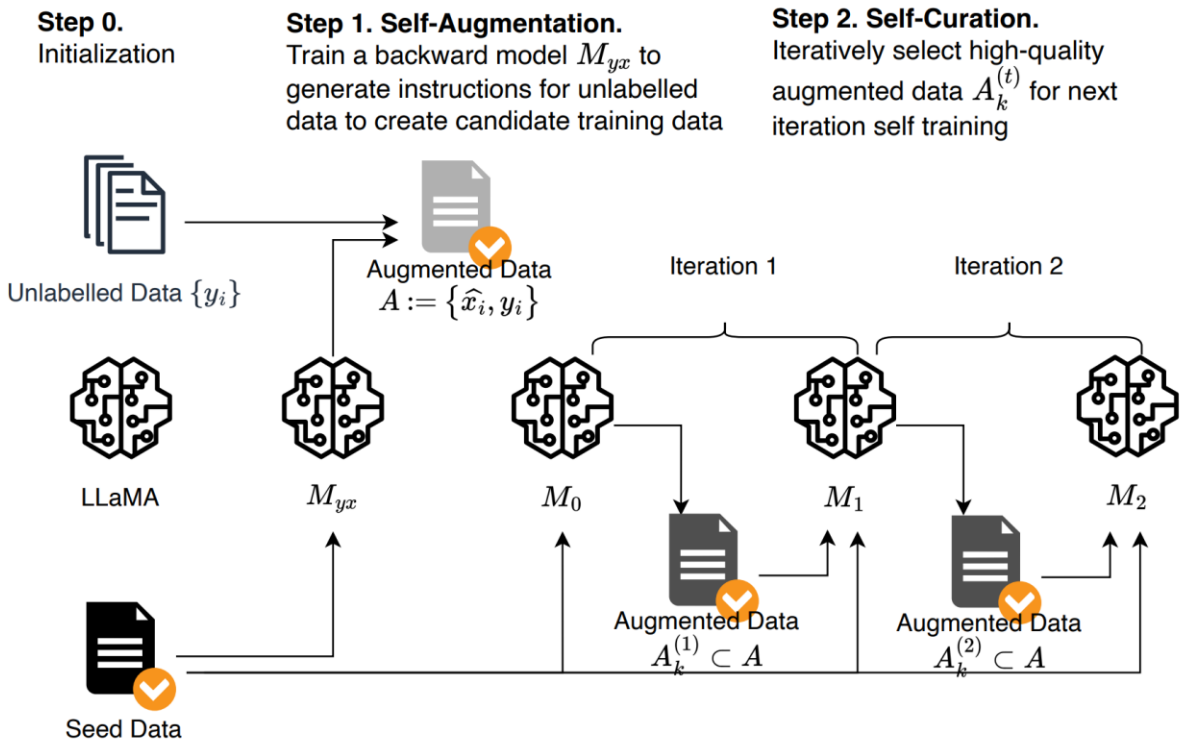
- Step 2: Reward Model Training
  - Instruction Reward Model (IRM)
  - Process-Supervised Reward Model (PRM)
- Step 3: Reinforcement Learning with PPO

WizardMath: Empowering Mathematical Reasoning for
Large Language Models via Reinforced Evol-Instruct (ICLR 2025 Oral)

Self-Alignment

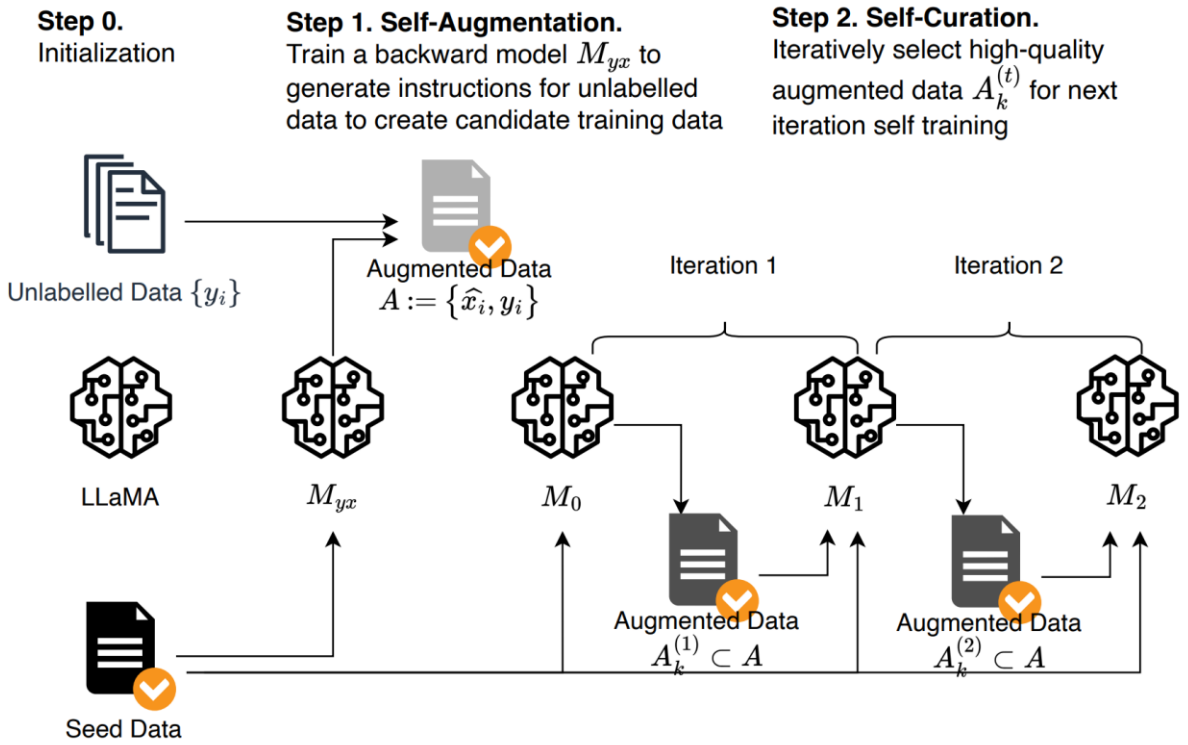
Self-Alignment with Instruction Backtranslation (ICLR 2024)

- Step 1: Initialization
  - A base LLaMA model
  - A small seed dataset of 3,200 high-quality human-annotated instruction-output pairs
  - A large corpus of unlabeled web text from Web



Self-Alignment

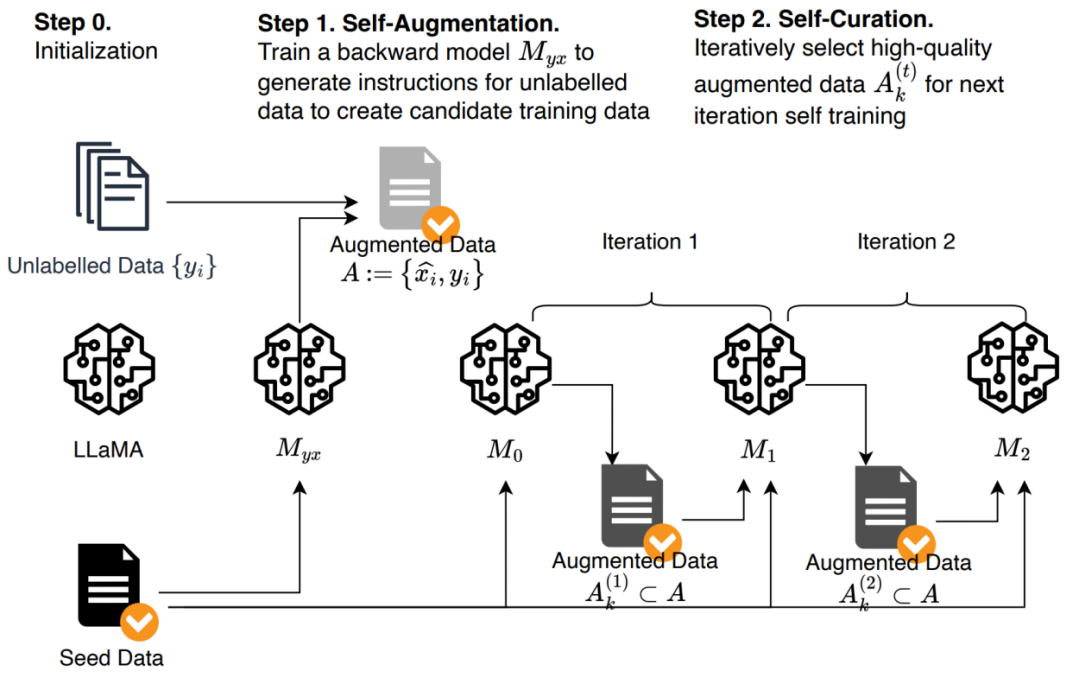
- Step 2: Self-Augmentation
  - Train backward model $M(y, x)$ on seed data (y, x)
  - Generate candidate instructions for documents
  - Create augmented dataset
  - Treat the document as the answer



Self-Alignment with Instruction Backtranslation (ICLR 2024)

Self-Alignment

- Step 3: Self-Curation (Iterative)
  - Train initial model M_0 on seed data only
  - Use M_0 to score quality of augmented pairs
  - Select high-quality pairs (score ⩾ 4 or 4.5) → A_1
  - Finetune M_0 on seed + A_1
  - Iterate: M_1 rescores data → A_2 → train M_2

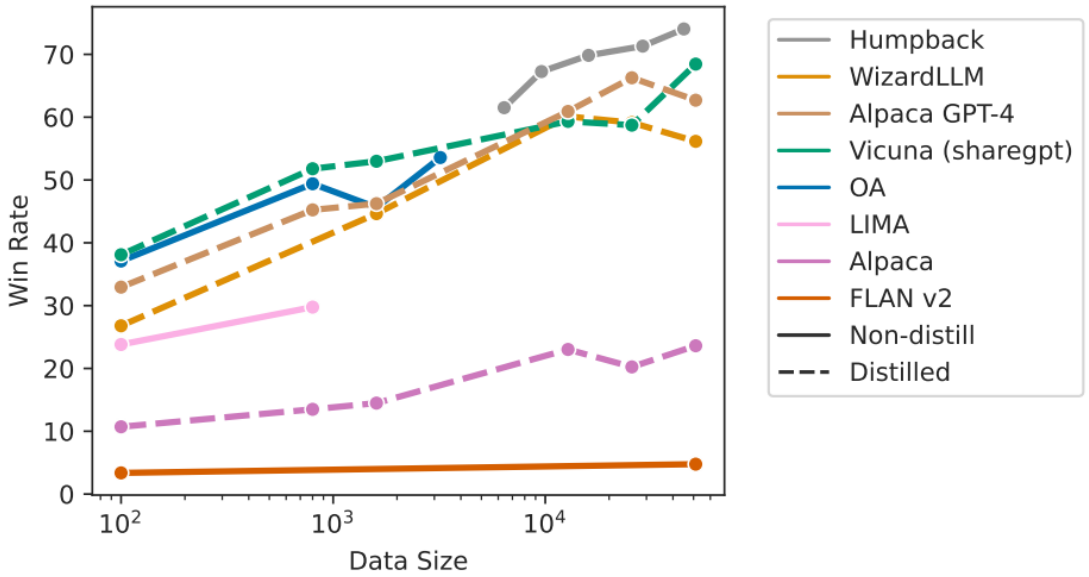


Self-Alignment with Instruction Backtranslation (ICLR 2024)

Alignment

- Advantages
  - Novel self-augmentation approach
  - Leveraging unlabeled data to ensure quality
  - Strong empirical results
- Disadvantages
  - Relies on seed data
  - May inherit and amplify biases from web data sources.
  - Quality scoring models achieve only 44-52% precision (Top Right Table), allowing many low-quality examples into training data.

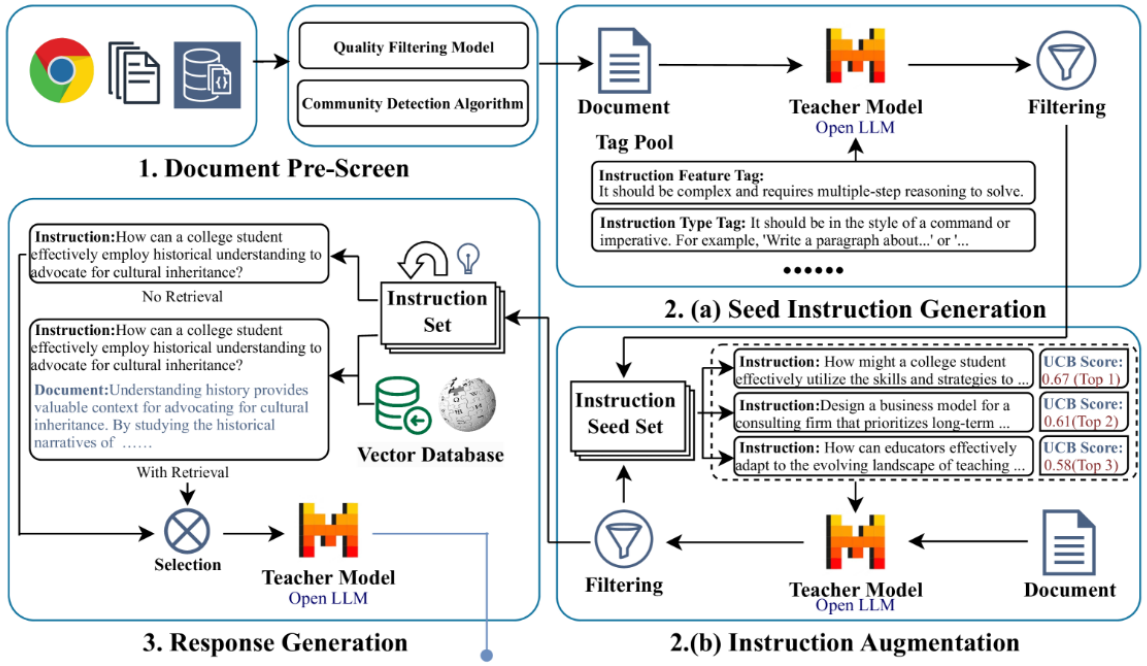
	Precision	Recall	Win Rate (%)
M_0	0.44	0.09	35.71 ±3.02
M_1	0.52	0.44	37.70 ±3.06
GPT-4	0.88	0.92	41.04±3.11



Self-Alignment with Instruction Backtranslation (ICLR 2024)

FANNO

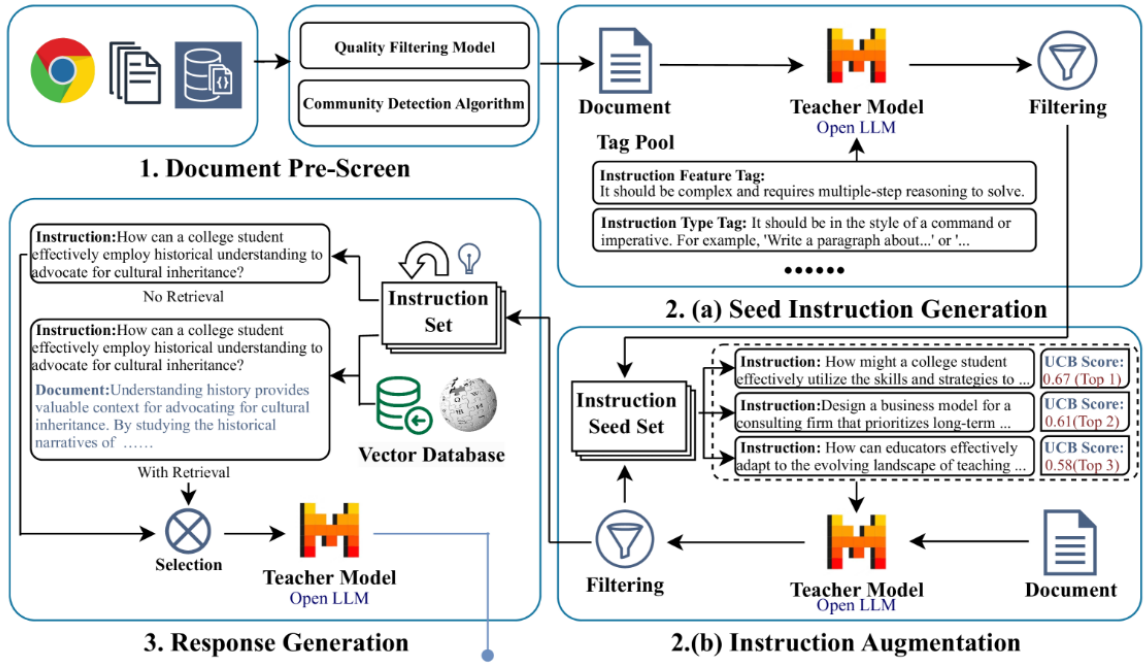
- The Problem of Instruction Data Creation
  - Human Annotation:
    - Expensive, time-consuming, and vulnerable to human cognitive biases
  - LLM-Generated Data (using proprietary models):
    - Reliance on these proprietary APIs introduces high costs and accessibility barriers.
- FANNO generates high-quality instruction data using only open-source LLMs.



FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only - ACL Anthology

FANNO

- The FANNO Framework
  - Document Pre-Screen: Filtering and preparing unlabeled text data
  - Instruction Generation: Creating diverse and complex instructions
  - Response Generation: Producing high-quality responses to the generated instructions
- The framework aims to address three key aspects of instruction data quality:
  - Diversity，Complexity，Faithfulness

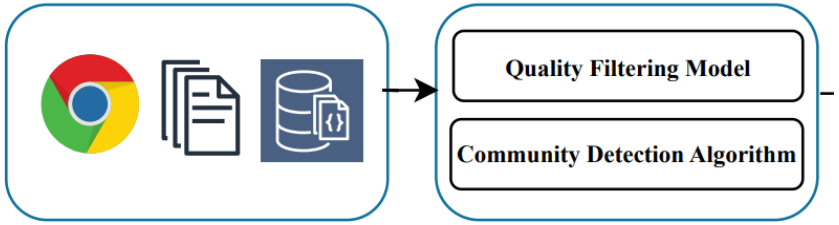


FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only (ACL 2025 Findings)

FANNO

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Document Pre-Screen
  - Segmentation and deduplication: Breaking text into segments and removing duplicates
  - Length-based filtering: Removing passages that are too short or too long
  - LLM-based filtering: Removing ambiguous content, privacy concerns, and advertisements
  - Community detection: Using a fast algorithm to cluster instruction embeddings and prioritize non-overlapping communities for diversity



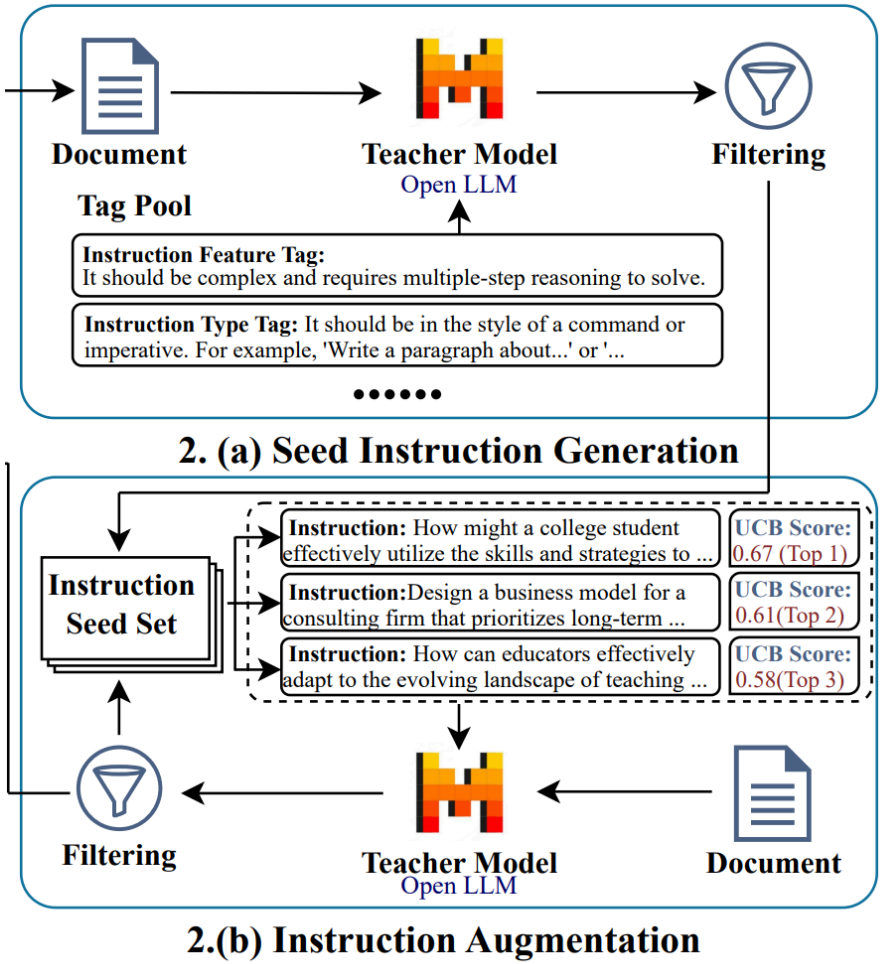
1. Document Pre-Screen

FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only (ACL 2025 Findings)

FANNO

- Instruction Generation
  - Seed Instruction Generation
    - Initial seed instructions are created by combining
      - Task type
      - Difficulty level tags
  - Instruction Augmentation
    - Think Different prompt template
    - Upper Confidence Bound (UCB) selection strategy

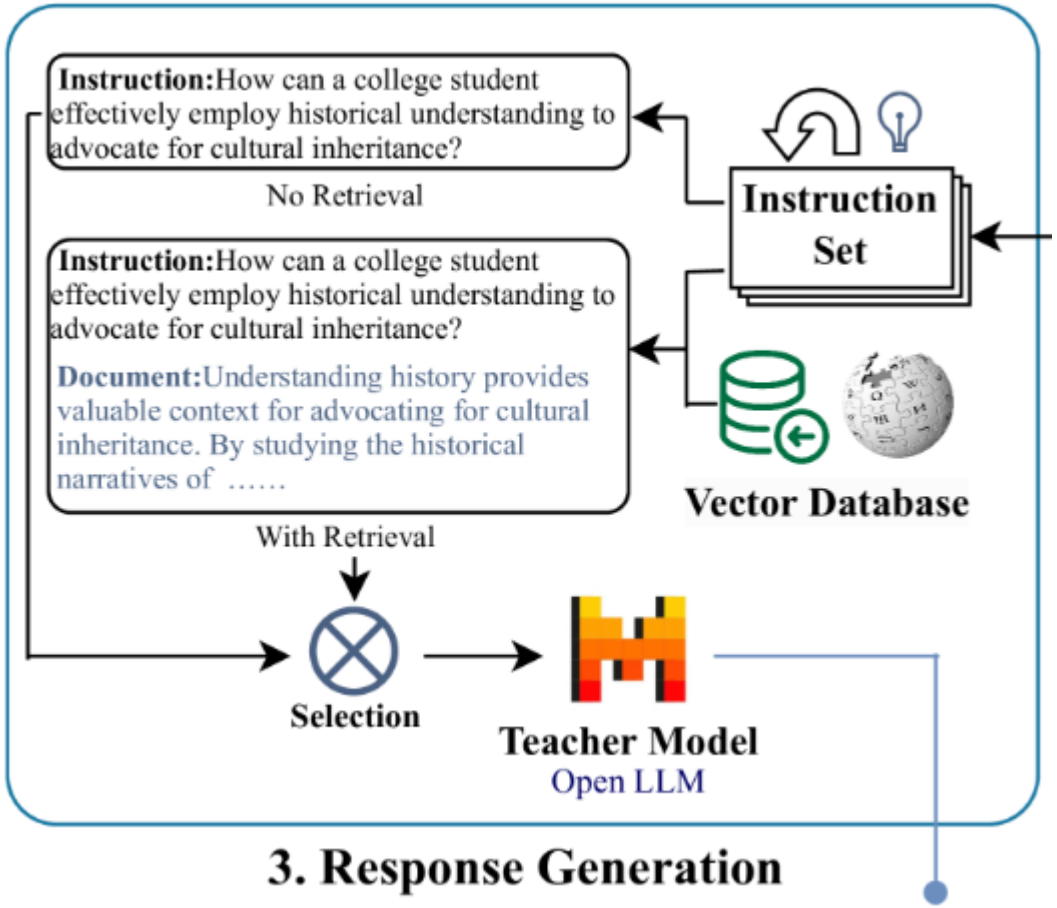
$$UCB(s) = \bar{x}_s + C\sqrt{\frac{2\ln N}{n_s}}.$$



FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only (ACL 2025 Findings)

FANNO

- Response Generation
  - Empty context: Generating responses directly from the instruction
  - Retrieval-Augmented Generation (RAG): Using relevant retrieved documents as context
- Advantage
  - Comprehensive and detailed
  - Directly addressing the instruction
  - Consistent and coherent

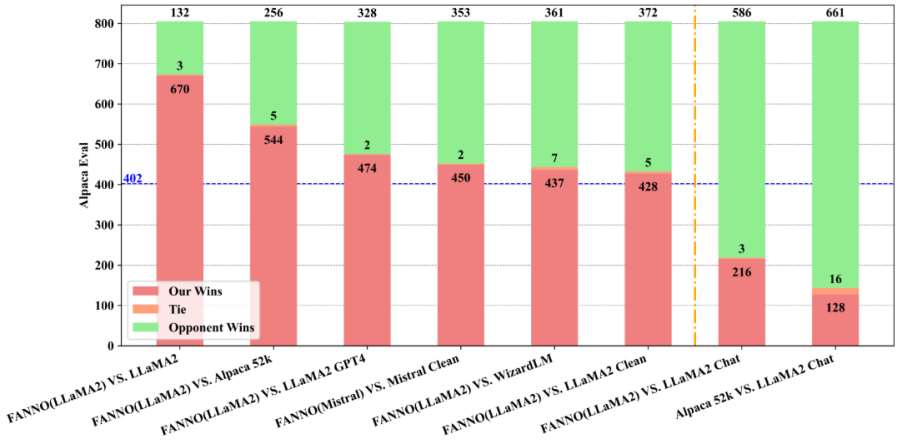


FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only (ACL 2025 Findings)

FANNO

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Models fine-tuned with FANNO-generated data showed exceptional performance across multiple benchmarks:



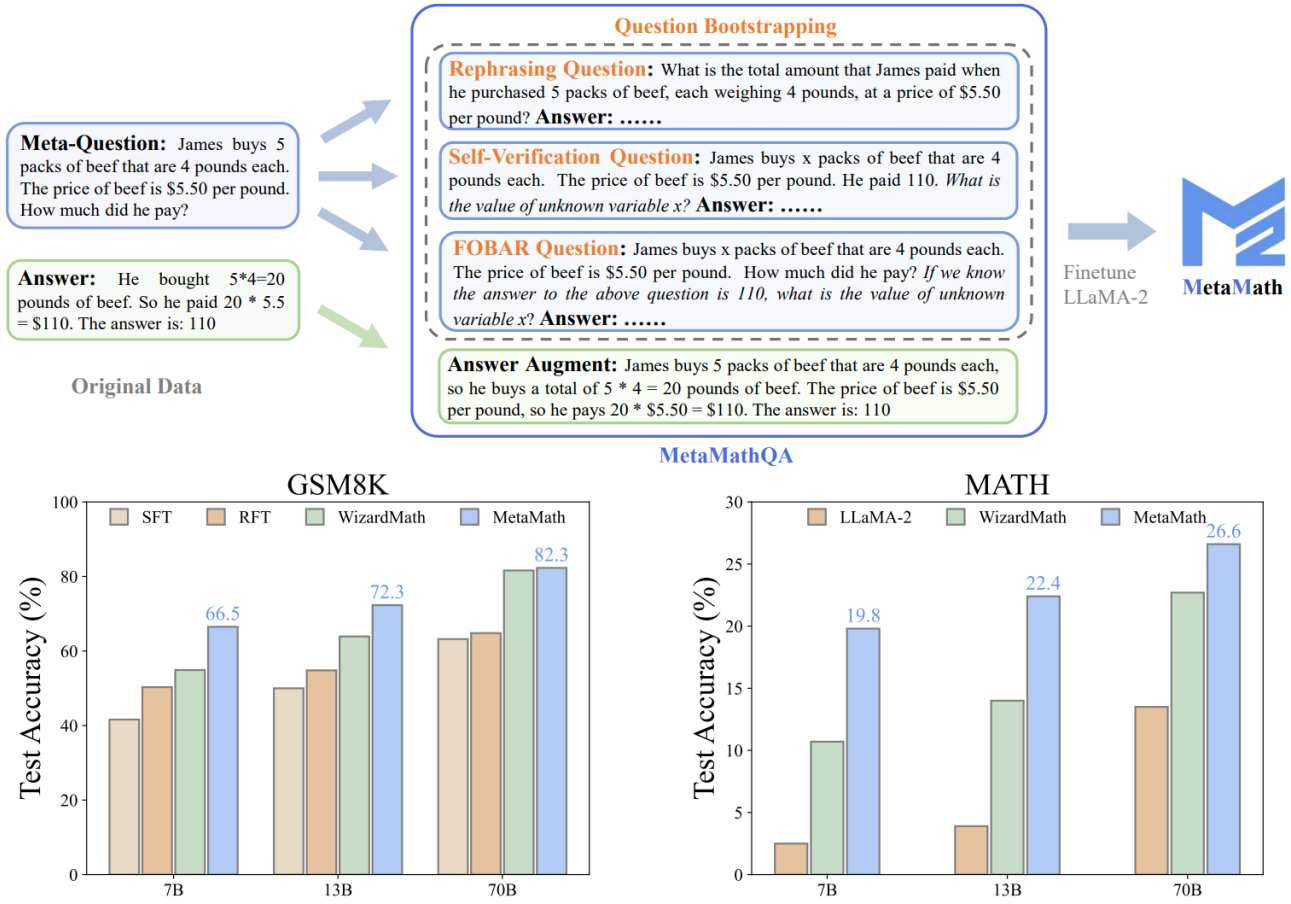
Model	Data Size	ARC	HellaSwag	MMLU	TruthfulQA	Average
Open-sourced Models based on LLaMA-2						
LLaMA-2-Base	–	54.10	78.71	45.80	38.96	50.76
LLaMA-2-Chat	–	54.10	78.65	45.69	44.59	55.76
LLaMA-2 + Alpaca-52k	52k	54.78	78.17	46.65	41.43	55.26
LLaMA-2 + Alpaca-GPT4	52k	56.66	78.78	46.96	51.02	58.35
LLaMA-2 + Alpaca-Cleaned	51.8k	56.40	80.16	47.02	50.53	58.53
LLaMA-2 + LIMA	1k	54.61	79.21	45.79	41.32	55.23
LLaMA-2 + WizardLM-70k	65k	54.01	78.66	45.61	38.99	54.32
LLaMA-2 + Muffin	68k	54.10	76.97	47.12	43.51	55.42
LLaMA-2 + FANNO	16k	55.63	79.45	46.84	51.01	58.23
Open-sourced Models based on Mistral-7B						
Mistral-7B-Instruct-v0.2	–	59.39	84.33	59.28	66.79	67.45
Mistral-7B-Base-v0.1	–	60.84	83.31	62.42	42.59	62.29
Mistral-7B-Base + Alpaca-GPT4	52k	63.65	82.18	59.29	43.98	62.29
Mistral-7B-Base + Alpaca-Cleaned	51.8K	64.51	83.68	59.76	52.00	64.99
Mistral-7B-Base + FANNO	16k	64.16	85.08	60.79	52.16	65.55

FANNO: Augmenting High-Quality Instruction Data with Open-Sourced LLMs Only (ACL 2025 Findings)

MetaMath

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

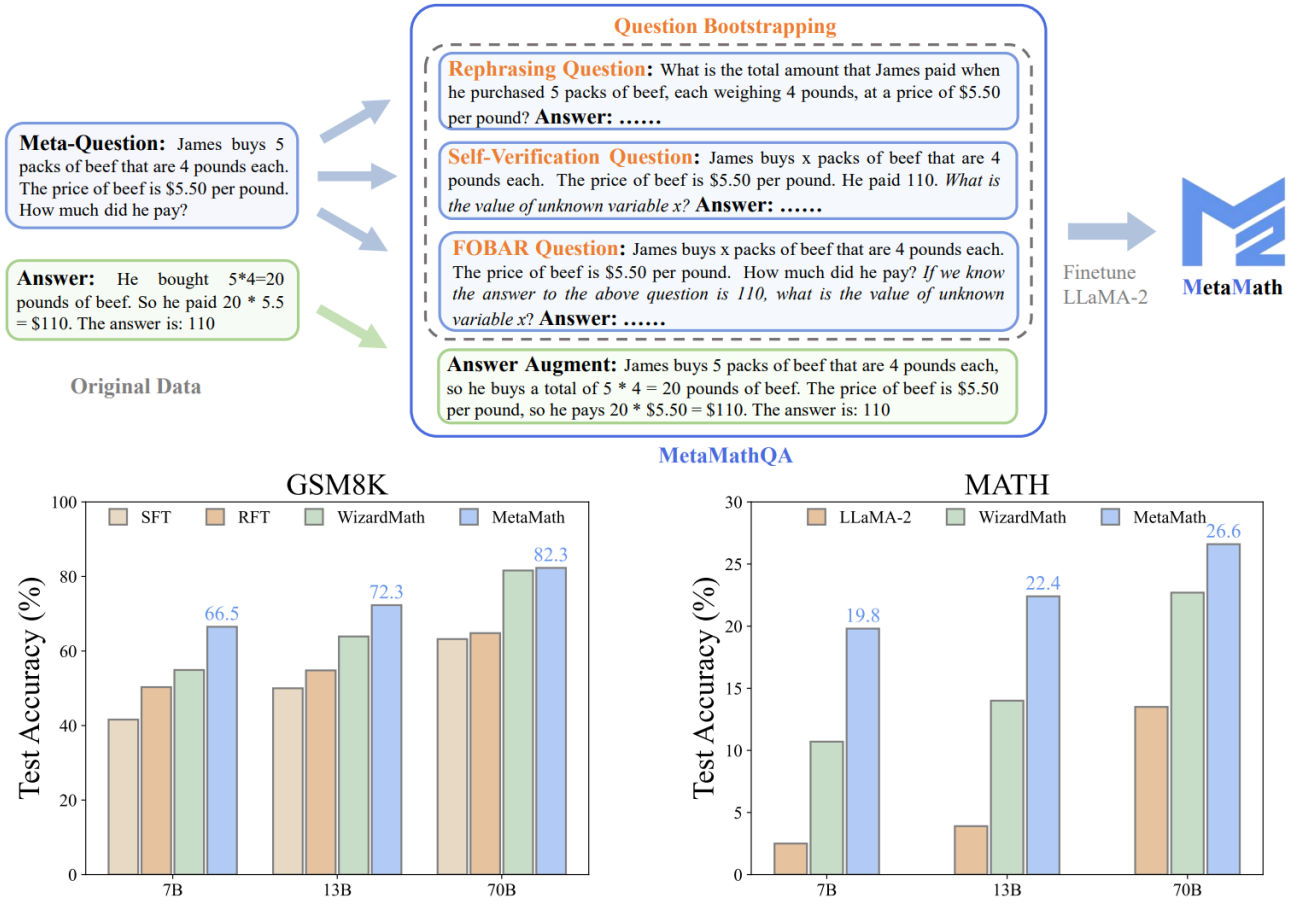
- MetaMathQA expands beyond traditional answer augmentation to generate diverse question formulations that capture the same mathematical concepts from multiple perspectives.



MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models (ICLR 2024 Spotlight)

MetaMath

- Answer Augmentation
  - generating multiple reasoning paths for existing questions using Chain-of-Thought prompting
  - For each original question q, the method produces multiple answer, retaining only those that yield correct answers to ensure data quality.



MetaMath

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

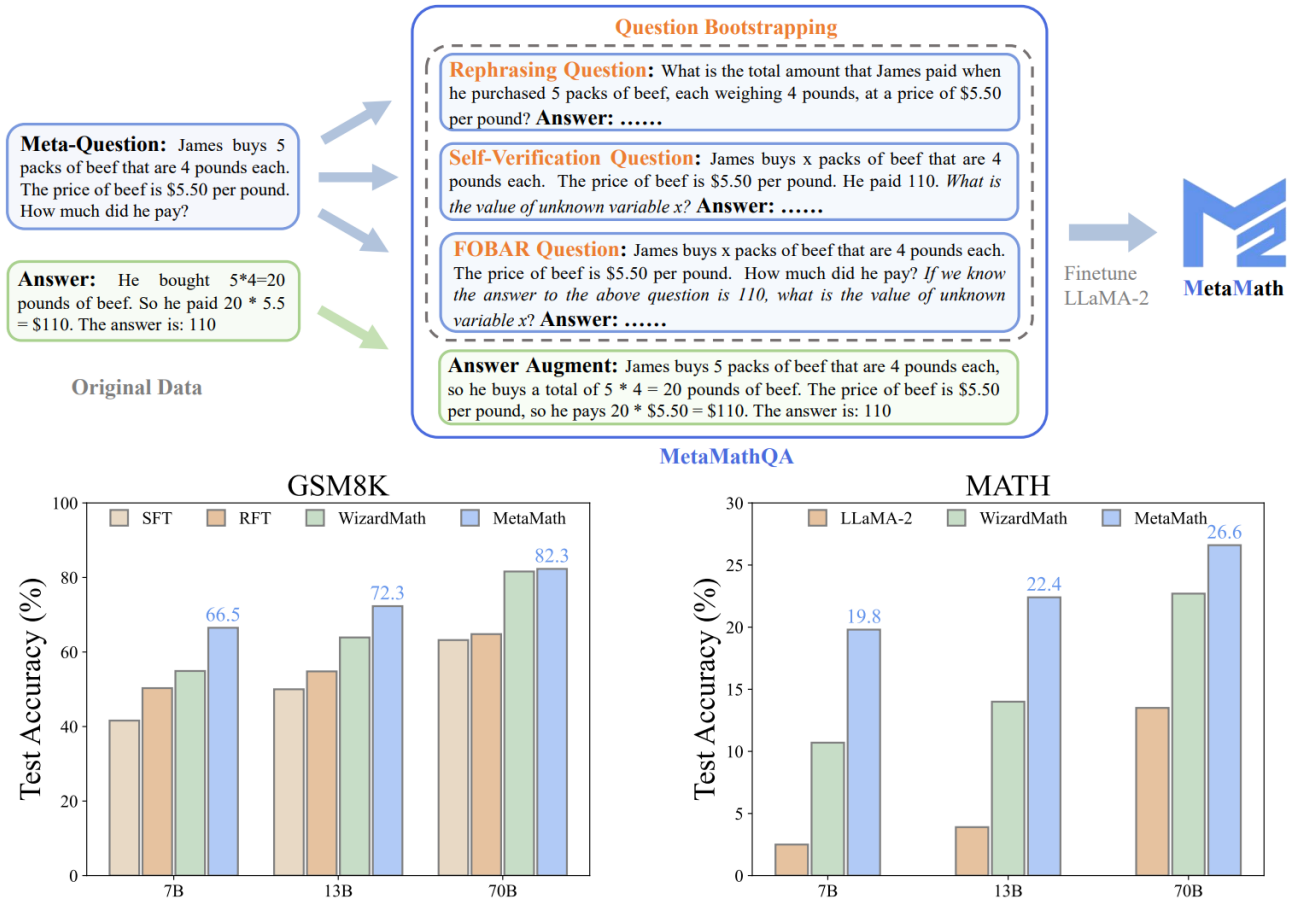
- LLM Rephrasing
  - GPT-3.5-Turbo generates syntactically and semantically varied reformulations of original questions.
  - Filter. Each rephrased question undergoes validation by generating its own reasoning path and answer, with only those matching the ground-truth answer being retained.
- Self-Verification (SV):
  - Transforms original questions into declarative statements
  - Poses new questions asking for masked variable values.
  - This approach enhances the model's ability to verify mathematical relationships and understand problem structures from different angles.

MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models

MetaMath

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- FOBAR (Forward-Backward Reasoning)
  - FOBAR generates backward reasoning questions by masking numerical values in original problems
  - Ask the model to deduce them given the final answer.
  - This directly addresses the "Reversal Curse" commonly observed in LLMs.

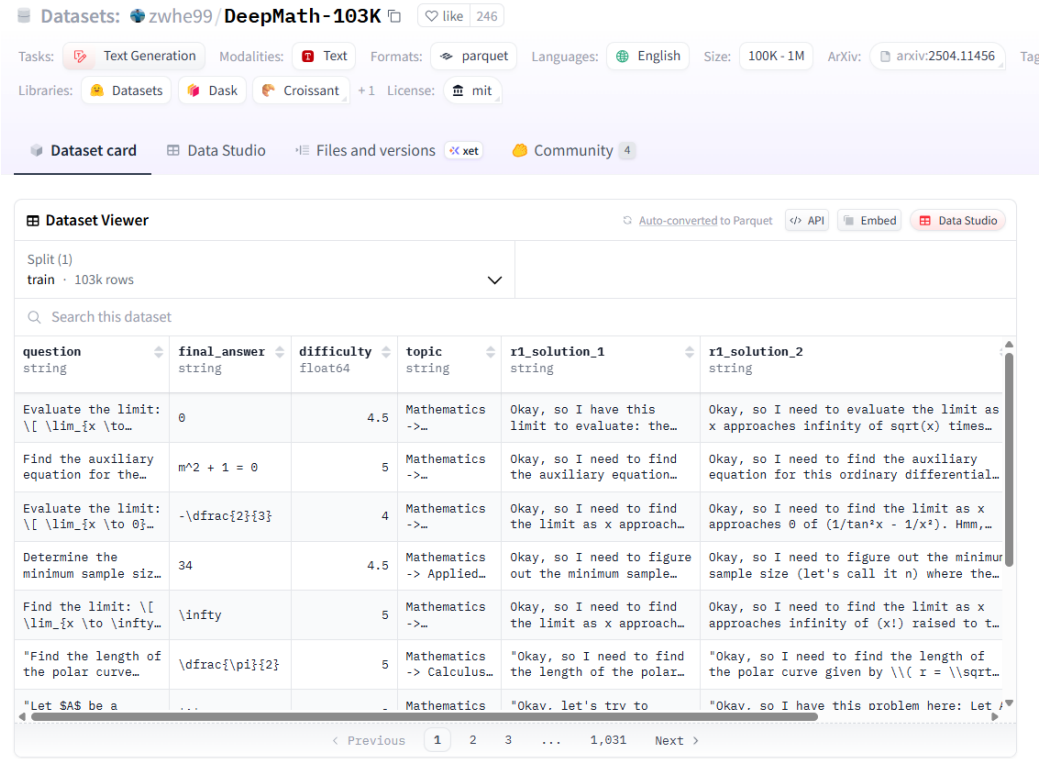


MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models

Math Data Synthesis

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- DeepMath-103K: A Large-Scale, Challenging, Decontaminated, and Verifiable Mathematical Dataset for Advancing Reasoning



Use DeepSeek-R1 for distillation to obtain high-quality response

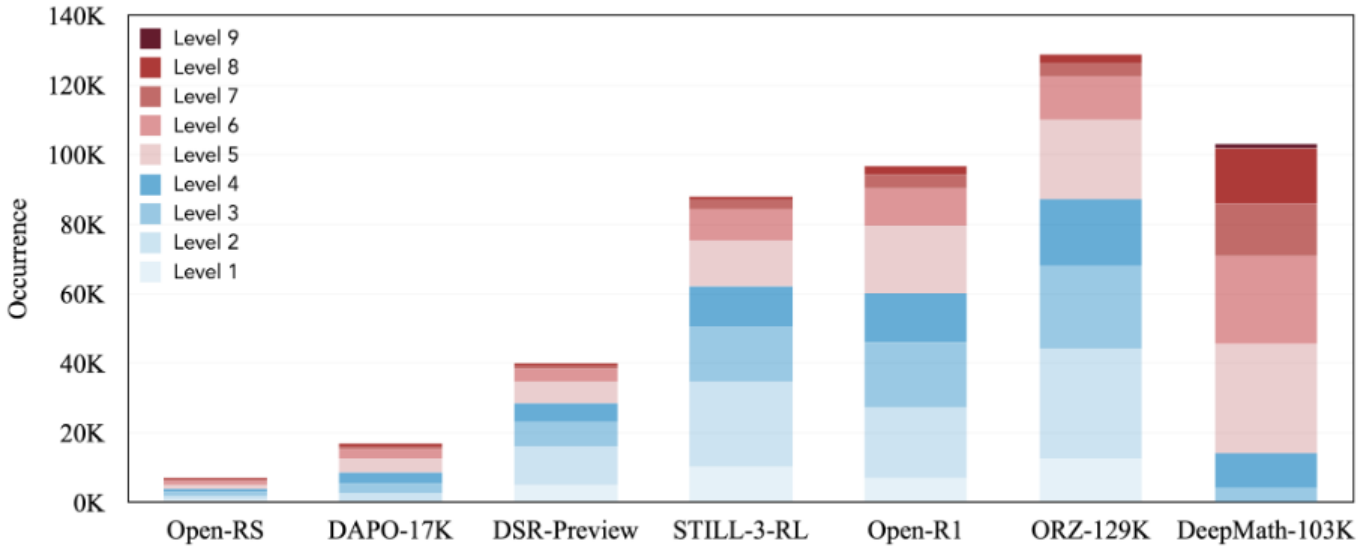
[2504.11456] DeepMath-103K: A Large-Scale, Challenging, Decontaminated, and Verifiable Mathematical Dataset for Advancing Reasoning

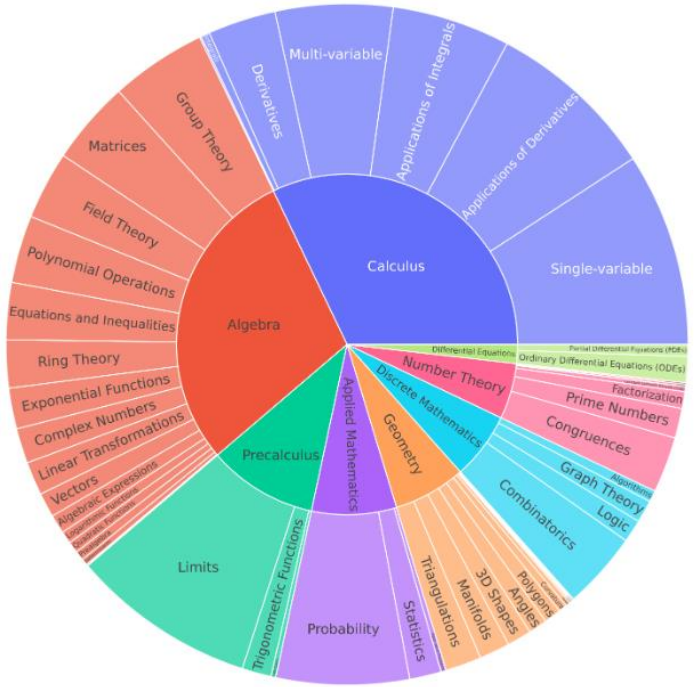
zwhe99/DeepMath-103K · Datasets at Hugging Face

Math Data Synthesis

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- DeepMath-103K: A Large-Scale, Challenging, Decontaminated, and Verifiable Mathematical Dataset for Advancing Reasoning



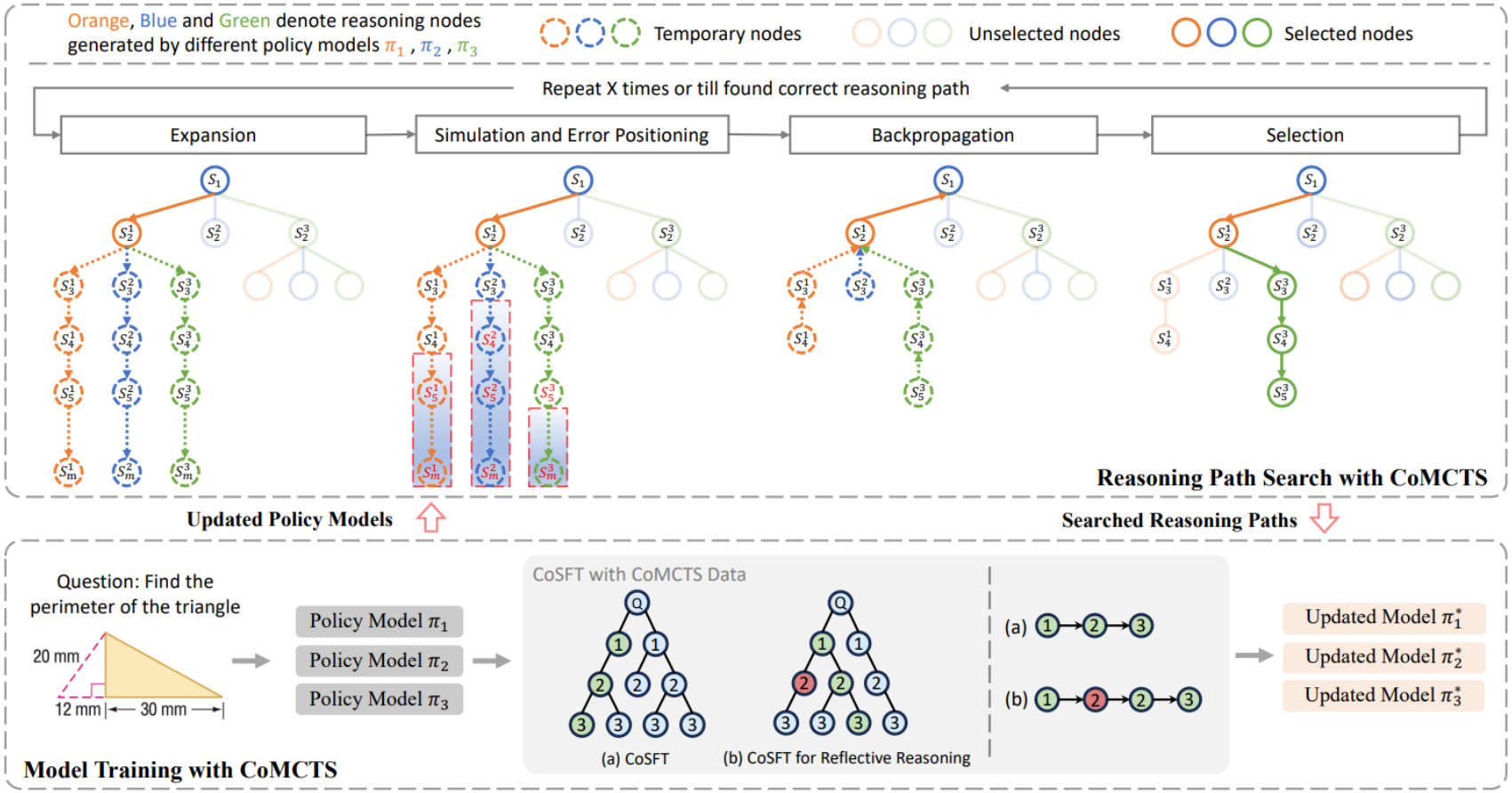


1. Challenging Problems: DeepMath-103K has a strong focus on difficult mathematical problems (primarily Levels 5-9), significantly raising the complexity bar compared to many existing open datasets.

2. Data Diversity and Novelty: DeepMath-103K spans a wide spectrum of mathematical subjects, including Algebra, Calculus, Number Theory, Geometry, Probability, and Discrete Mathematics.

Multimodality

- Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search (NeurIPS 2025 Spotlight)



Multimodality

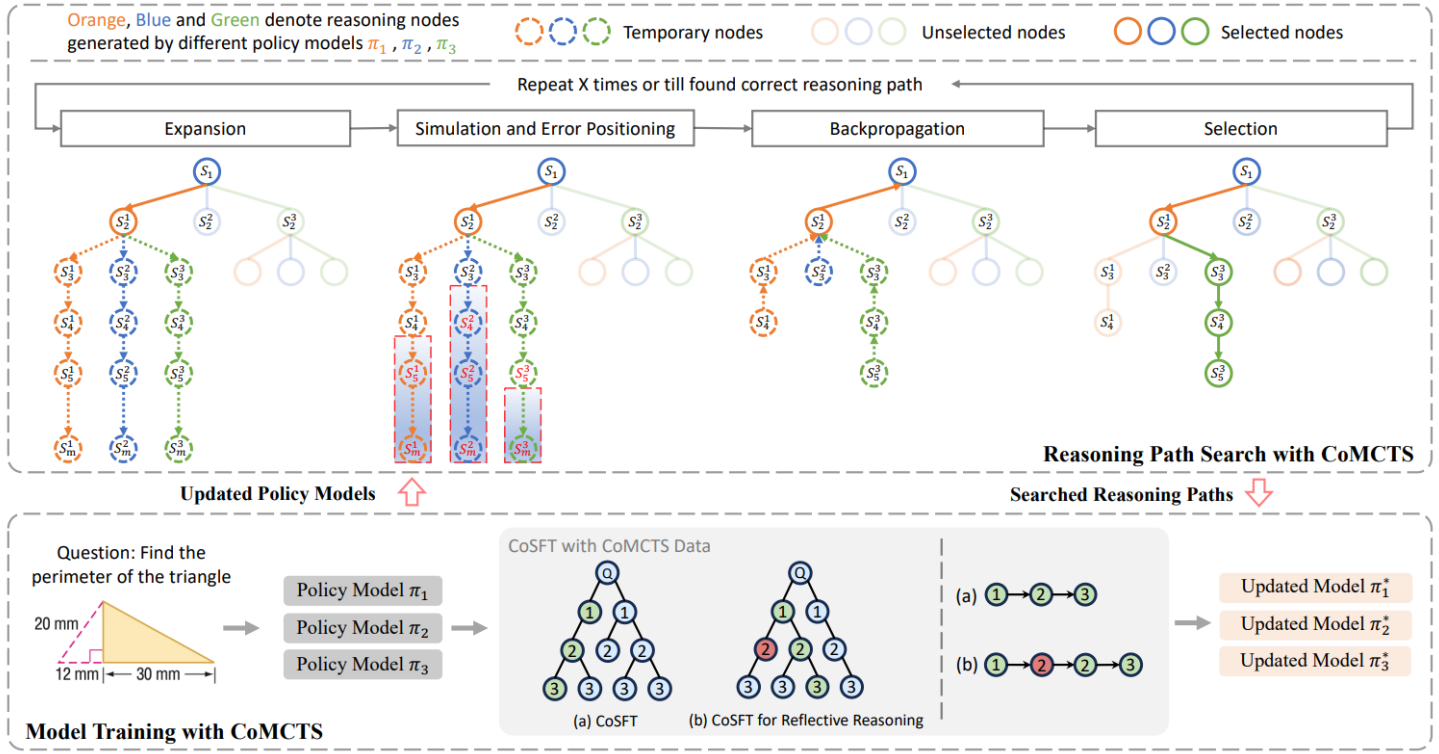
南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Background and Motivation

- Applying MCTS to MLLMs faces two challenges:
  - Search Effectiveness: MLLMs tend to get trapped in "homogeneous low-quality nodes" when self-bootstrapping, leading to poor exploration of reasoning spaces and low success rates.
  - Search Efficiency: Traditional MCTS explores one node at a time, making it computationally prohibitive for resource-intensive MLLMs that require massive iterations.

Multimodality

- Collective Monte Carlo Tree Search (CoMCTS) represents a fundamental departure from traditional MCTS by leveraging knowledge from a collective group of MLLMs to collaboratively search for effective reasoning paths.
  - Expansion
  - Simulation and Error Positioning
  - Backpropagation
  - Selection
  - Reflective Reasoning Extension

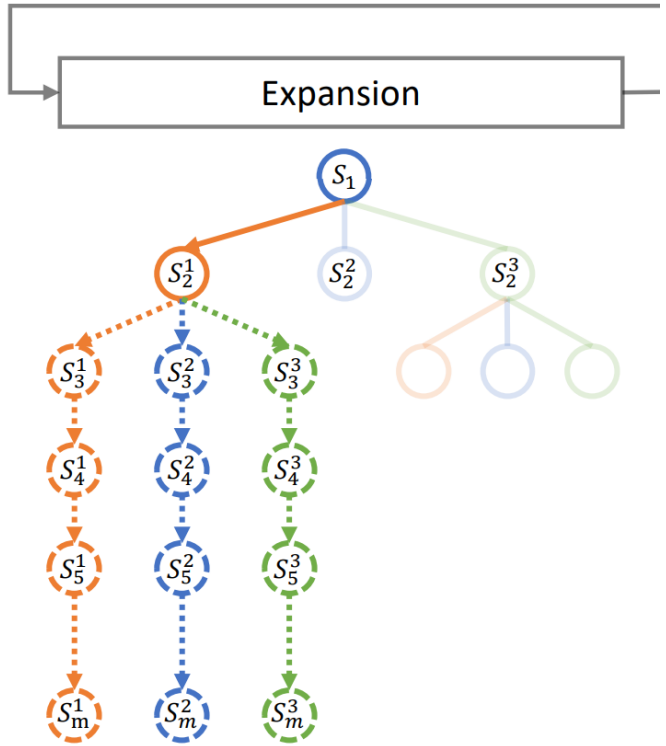


Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- Expansion
  - CoMCTS performs collective expansion where each MLLM in the group generates potential reasoning paths in parallel from the current leaf node.
  - Advantage
    - Diverse candidate reasoning nodes
    - Prevents the search from getting stuck in low-quality reasoning spaces of individual models
    - Significantly boosts search efficiency by exploring multiple paths simultaneously

Orange, Blue and Green denote reasoning nodes generated by different policy models π_1 , π_2 , π_3



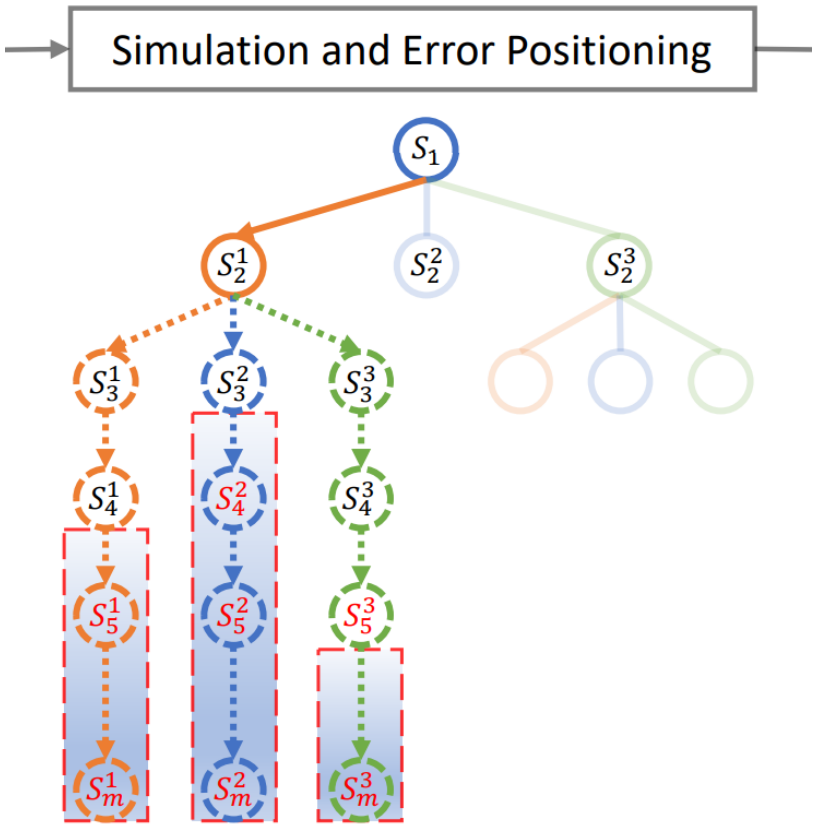
Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- Simulation and Error Positioning
  - Let multiple models collectively evaluate candidate inference nodes (using prompt), and those below the threshold are considered as error nodes and pruned.

$$R(s_i^j) = \frac{1}{K}\sum_{l=1}^{K}\pi_l(\cdot|\text{prompt}_{\text{eval}}, Q, \text{Parent}(s_i^j), s_i^j)$$

$$S^*_{\text{candidate}} = \{s_i^j \in S_{\text{candidate}} | R(s_i^j) >= t\}$$



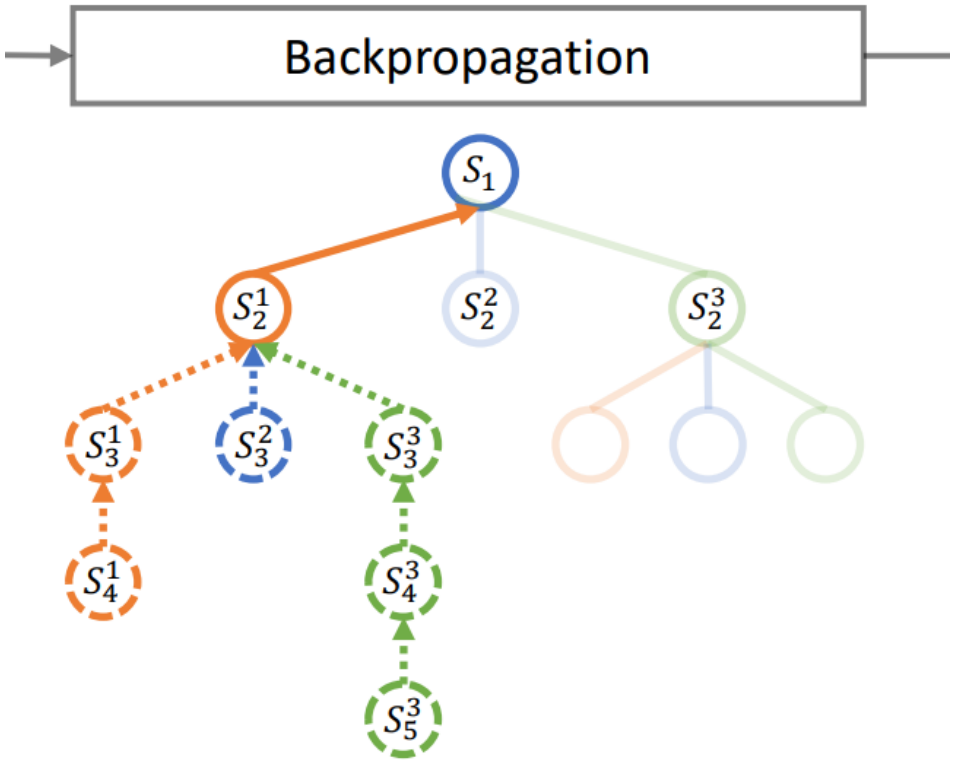
Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- Backpropagation
  - Update visit count N(s) and node value V(s) for all nodes along the newly expanded path.

$$V(s) \leftarrow \frac{N(s) \cdot V(s) + \sum_{s_l \in \text{Child}(s)} R(s_l)}{N(s) + \text{CountChild}(S^*_{\text{candidate}}, s)},$$

$$N(s) \leftarrow N(s) + \text{CountChild}(S^*_{\text{candidate}}, s),$$

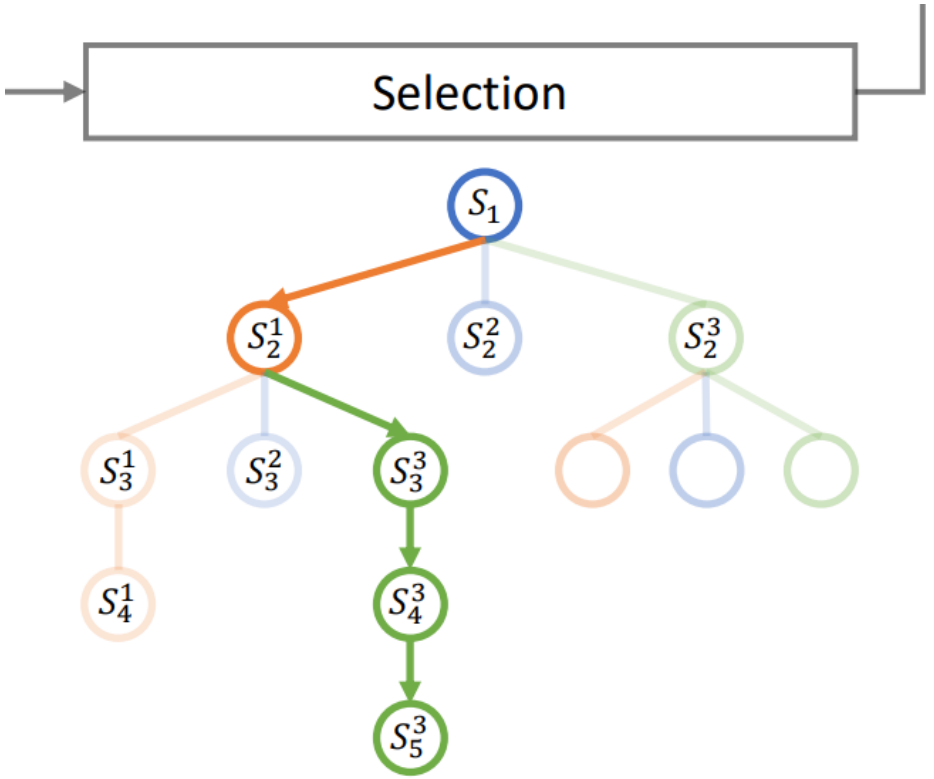


Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- Selection
  - Use Upper Confidence Bound (UCB) to balance exploration and exploitation.

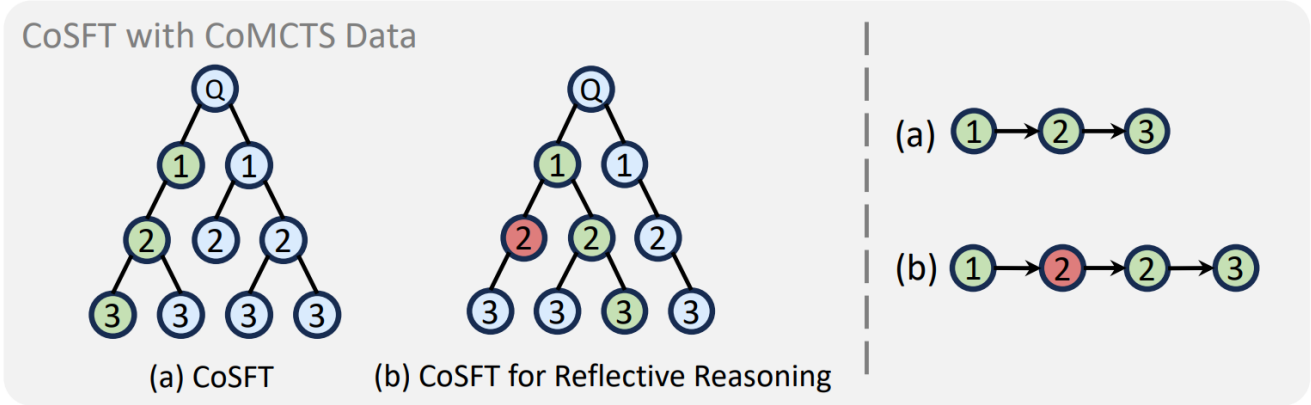
$$s_m^{k^*} = \arg\max_{s \in S^*_{\text{candidate}}} V(s) + c \cdot \sqrt{\frac{\log N(\hat{s})}{1 + N(s)}}$$



Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- CoMCTS for Reflective Reasoning
  - Negative Sibling Identification: The algorithm identifies "negative sibling nodes" for nodes on effective reasoning paths - these are nodes at the same hierarchical level with significantly lower UCB values
  - Reflective Path Construction: Reflective trajectories are formed by concatenating negative sibling nodes, reflection prompts (e.g., "The previous reasoning step is wrong and let's rethink it again"), and corresponding positive nodes



Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

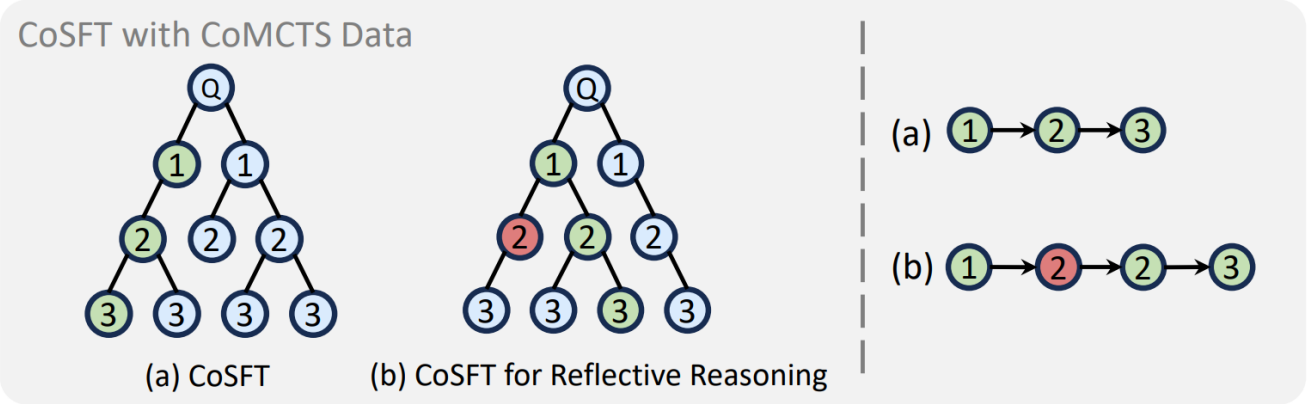
- Training Strategy: CoSFT
  - Standard CoSFT:

$$\mathcal{L}_{\text{CoSFT}}(\pi_k) = \sum_{(Q,Y)\in\mathcal{D}} \log \pi_k(Y|Q),$$

  - CoSFT for Reflective Reasoning

$$\mathcal{L}_{\text{CoSFT-Re}}(\pi_k) = \sum_{(Q,Y_{\text{reflect}})\in\mathcal{D}} \log \pi_k(Y_{\text{reflect}}|Q),$$

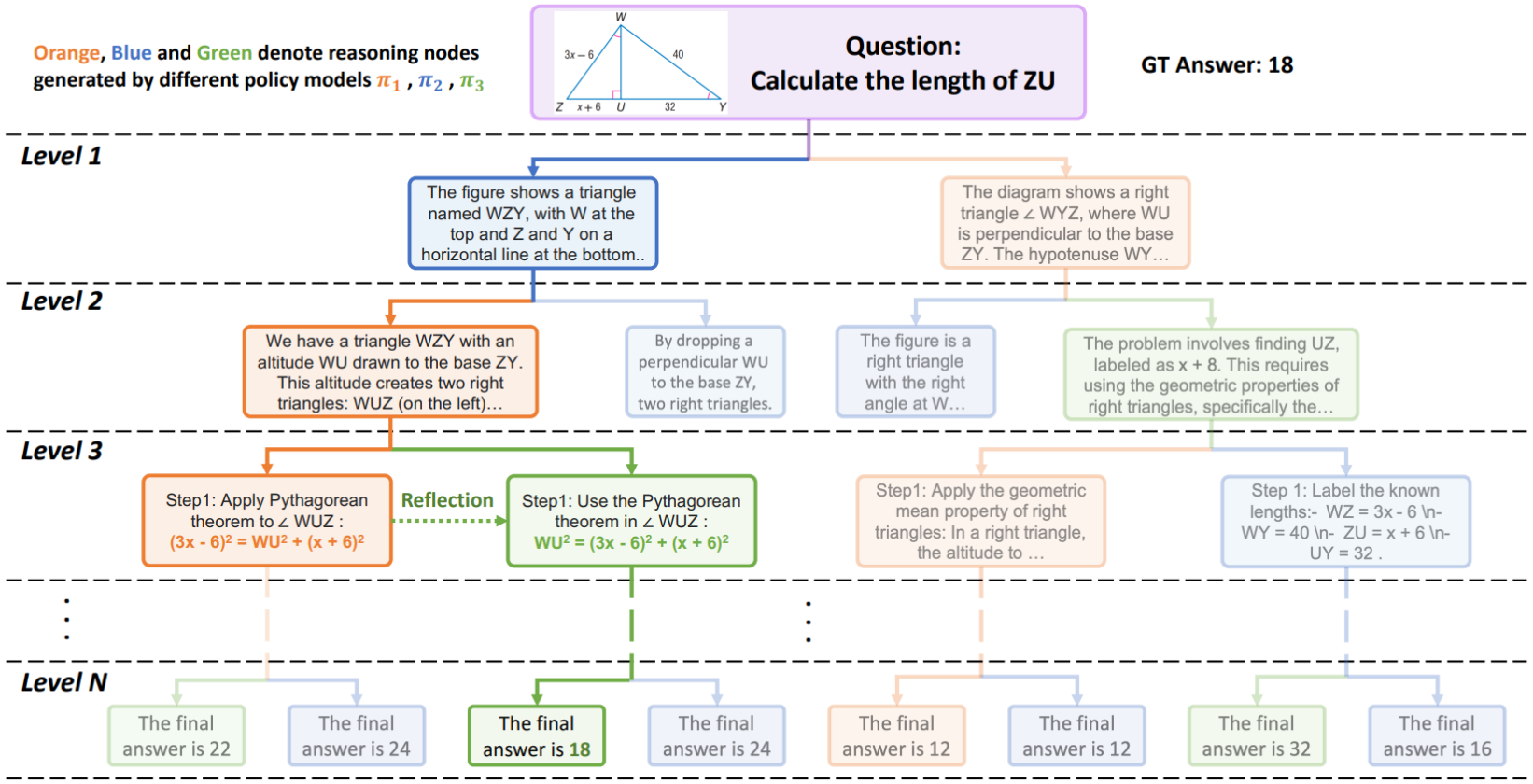
  - Total 260K Data (*Reflection Data 15K*)



Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

Multimodality

- Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

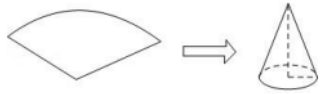


Multimodality

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Sea

Mulberry creates rich, explicit and well-defined reasoning steps with comprehensive understanding, ultimately arriving at the correct answer.



Question: Use a sector paper sheet with a central angle of 120.0 and a radius of 6.0 to roll into a conical bottomless paper cap (as shown in the picture), then the bottom perimeter of the paper cap is ()

Choices:
(A) 2 π cm; (B) 3 π cm;
(C) 4 π cm; (D) 5 π cm

Correct Answer: (C)

LLaVA: We need to find the derivative of f(x) at x=2 and compare it with the derivative of f(x) at x=5. We can do this by using the power rule of differentiation, which states that if f(x) = x^n, then f'(x) = nx^(n-1). At x=2, f'(2) = 2^(n-1) = 2^3 = 8. At x=5, f'(5) = 5^(n-1) = 5^3 = 125. Therefore, the derivative of f(x) at x=2 is larger than that at x=5. The correct answer is (A).

Qwen2-VL: First, let's find the circumference of the base of the cone. The circumference of the base is equal to the length of the arc of the sector paper sheet.

The formula for the circumference of a circle is $C = 2 \pi r$, where **r** is the radius. In this case, the radius of the base of the cone is equal to the radius of the sector paper sheet, which is 6.0 cm.

So, the circumference of the base of the cone is: $C = 2 \pi \times 6.0 = 12 \pi$ cm.

However, the question asks for the bottom perimeter of the paper cap, which is the circumference of the base of the cone. Therefore, the bottom perimeter of the paper cap is:12π cm.

So, the correct answer is D

Mulberry:

Image Description: A diagram shows a circular sector being transformed into a cone without its base. The sector has a central angle labeled as **120°** and a given radius length of **6**.

Rationales: To find the circumference or "bottom perimeter" of the resulting cone from rolling up this sector, we need to understand that when you form a cone using such a sector, the arc length becomes the circumference of the base circle of the cone.

Let's think step by step.

Step 1: Calculate the arc length of the sector which forms the circumference of the cone's base. Arc Length = central angle/360 × 2πr. Here, r=6 cm and central angle = 120°.

Step 2: Substitute values into the formula:
Arc Length = (120/360) × 2π × 6.
Simplify it further to get the actual arc length.

Step 3: Perform calculations: Arc Length = 1/3 × 12π = 4π. This means the circumference of the cone's base equals **4π**.

The final answer is: C

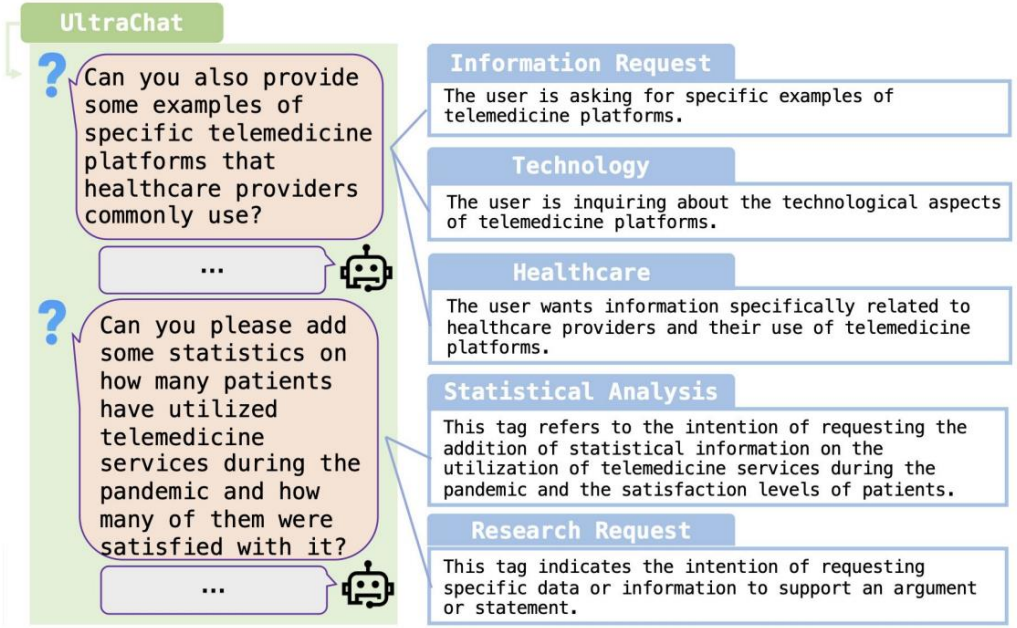
Approaches to Data Filtering

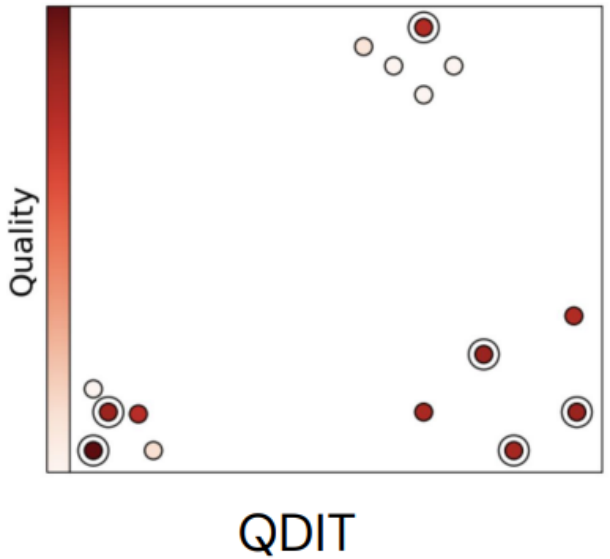
- Diversity filtering
- Quality filtering
- Correctness filtering

https://synth-data-acl.github.io/static/slides/slides.pdf

Diversity Filtering

- Filter similar examples as defined by
  - Rouge-L (Self-Instruct; Impossible Distillation)
  - Embedding similarity (QDIT, DiverseEvol, DEITA)
  - Semantic tags (#InsTag)

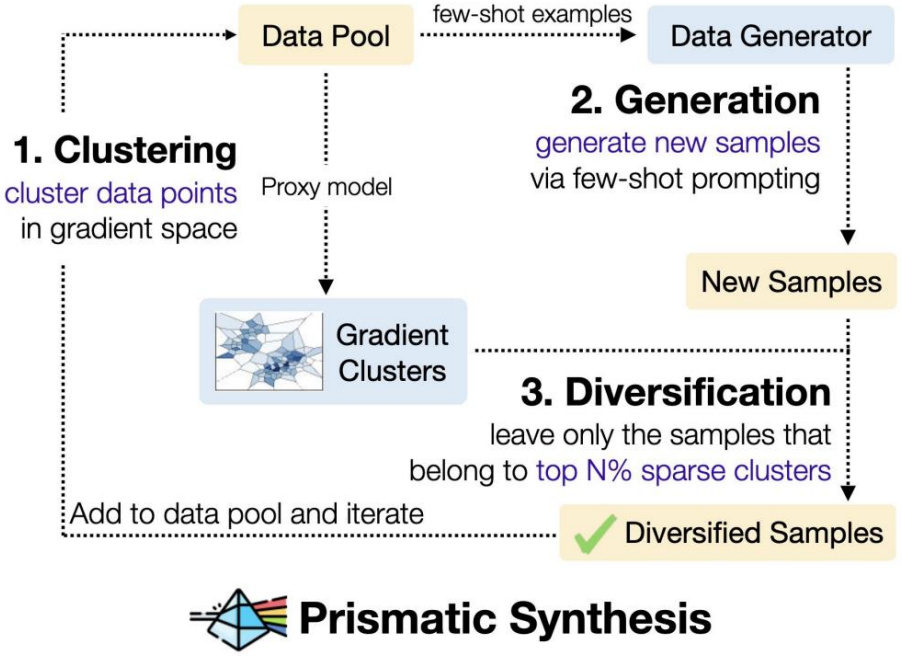


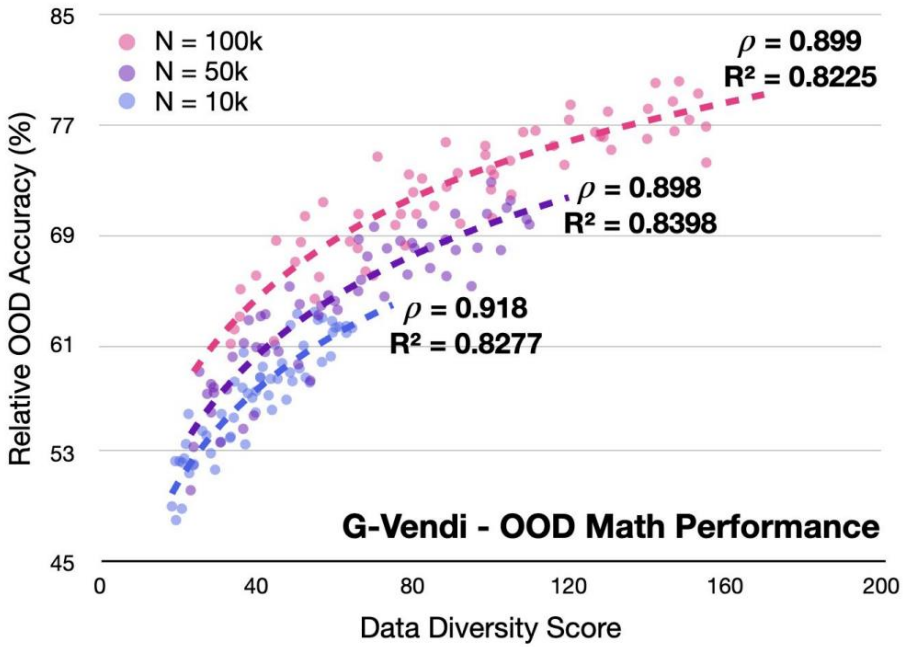


https://synth-data-acl.github.io/static/slides/slides.pdf

Diversity Filtering: Gradients

- Measure diversity of data in loss gradients
- Higher data diversity ⇒ more robust models

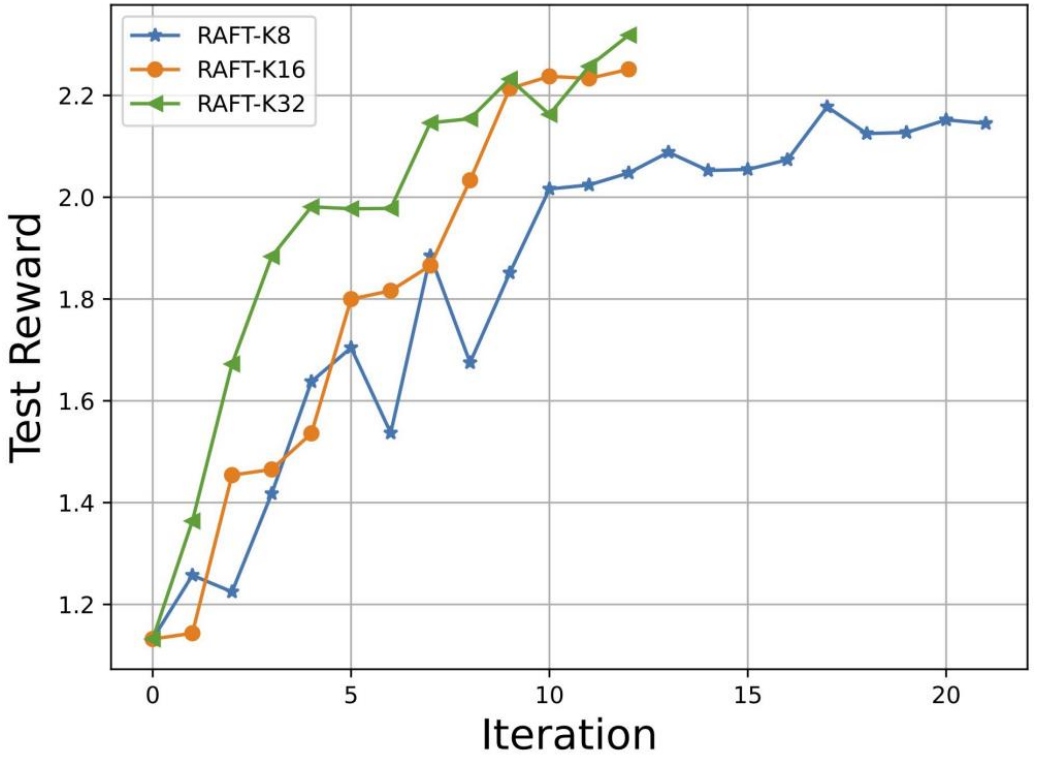




Prismatic Synthesis: Gradient-based Data Diversification Boosts Generalization in LLM Reasoning (Jung et al., 2025)
https://synth-data-acl.github.io/static/slides/slides.pdf

Quality Filtering: Reward Models

- Sample K responses and take the one with the highest reward, then SFT on the best-of-K responses

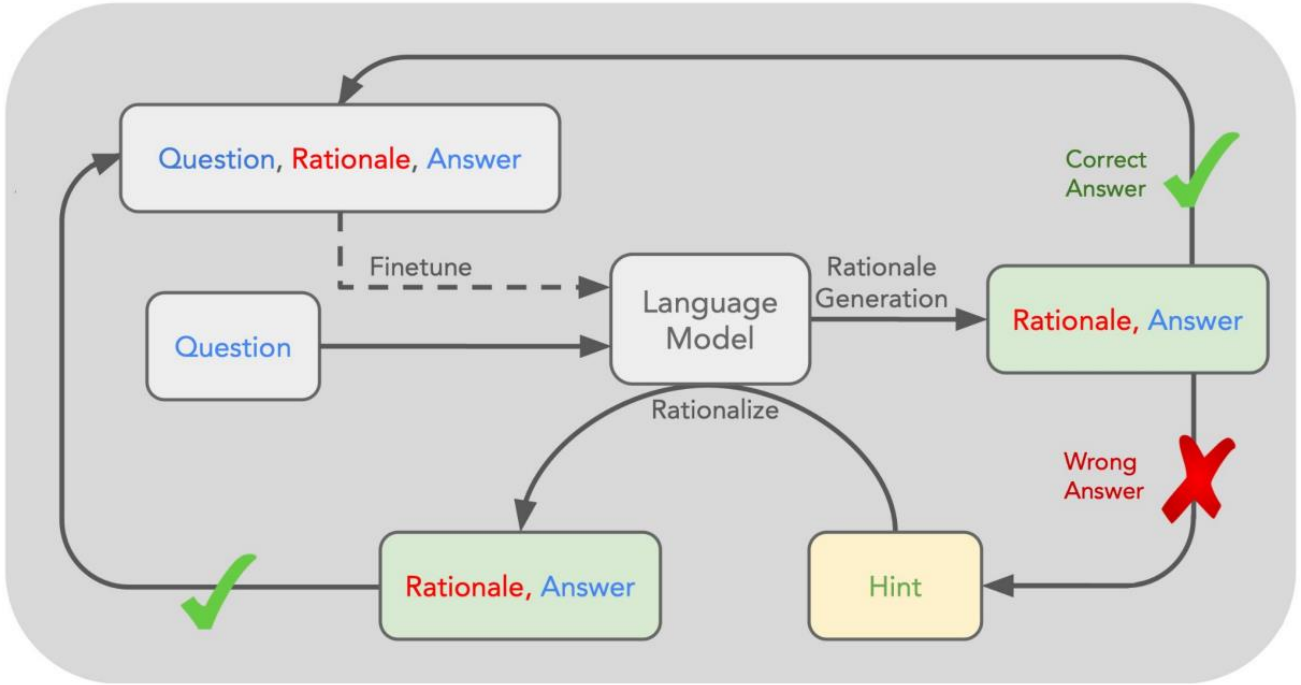


RAFT: Reward rAnked FineTuning for Generative Foundation Model Alignment

https://synth-data-acl.github.io/static/slides/slides.pdf

Correctness Filtering: Final Answer Verification

- When generating synthetic reasoning data, only keep generations whose final answers are correct




```
Q: What can be used
to carry a small dog?
Answer Choices:
(a) swimming pool
(b) basket
(c) dog show
(d) backyard
(e) own home
A: The answer must be
something that can be
used to carry a small
dog. Baskets are
designed to hold things.
Therefore, the answer
is basket (b).
```

STaR: Bootstrapping Reasoning With Reasoning
https://synth-data-acl.github.io/static/slides/slides.pdf

Challenges and Limitations of Synthetic Data

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Misuse of synthetic data might proliferate misinformation
- Synthetic data might cause ambiguity in AI alignment
- Training with synthetic data makes evaluation decontamination harder

Best Practices and Lessons Learned on Synthetic Data

Directions for Future Work

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Synthetic data scaling
- Further improving quality and diversity of synthetic data
- Towards high-fidelity and more efficient scalable oversight
- The emergent self-improvement capability

Thank you