SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection

Liangxin Liu[1], Xuebo Liu[1], Derek F. Wong[2], Dongfang Li[1], Ziyi Wang[1], Baotian Hu[1], Min Zhang[1]

[1] Institute of Computing and Intelligence, Harbin Institute of Technology, Shenzhen, China
[2] NLP²CT Lab, Department of Computer and Information Science, University of Macau

Agenda

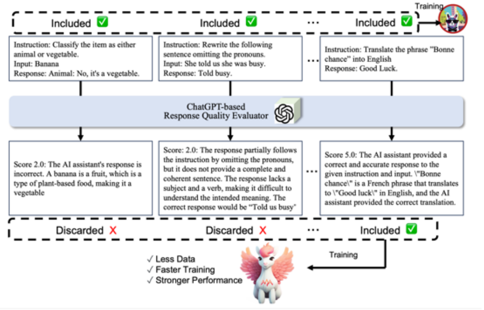
1. Introduction
2. SelectIT
3. Experiment
4. Analysis
5. Conclusion

1 Introduction

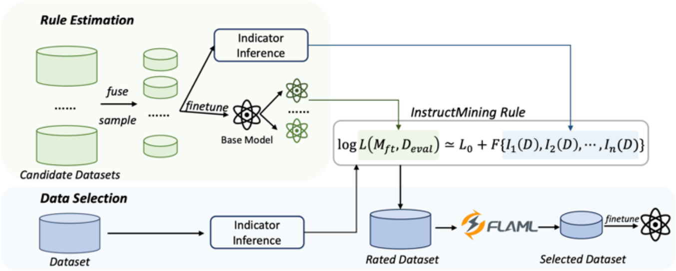
Introduction

Data Selection Definition

- Instruction tuning (IT) is crucial for improving large language model (LLM) interactions.
- However, high-quality data selection often relies on external resources, restricting broader application.
  - Some researchers use closed-source LLMs or additional datasets to evaluate and train models for optimized IT data.



ALPAGASUS[1]



Instruction Mining[2]

[1] AlpaGasus: Training A Better Alpaca with Fewer Data. ICLR, 2024

[2] Instruction Mining: Instruction Data Selection for Tuning Large Language Models. arXiv, 2023

Introduction

Problem

Our objective is to explore the question:

- Exiting advanced data selection strategies rely heavily on external models or data.
- How can we overcome the existing limitations so that can select data efficiently.

✓ We propose SelectIT, a novel IT data selection method which exploits the uncertainty of LLMs without using additional resources.

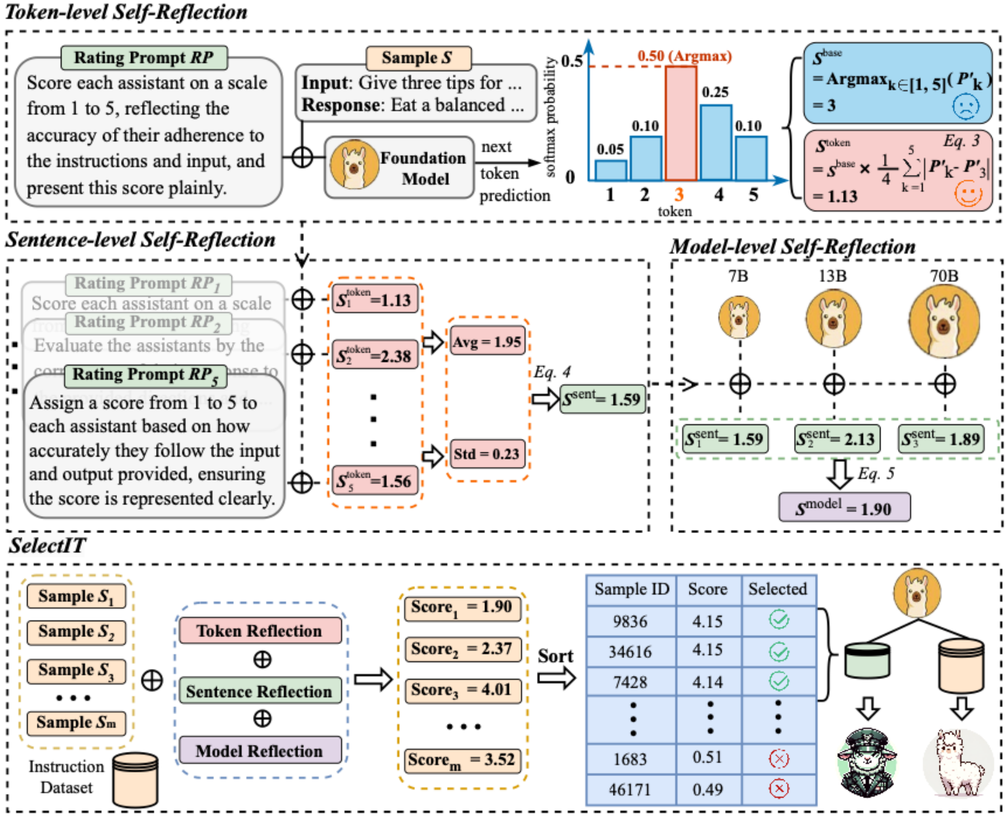
✓ SelectIT can substantially improve the performance of LLMs across a variety of foundation models and domain-specific tasks.

✓ Our analysis suggests that longer and more computationally intensive IT data may be more effective, offering a new perspective on the characteristics of optimal IT data.

2

SelectIT

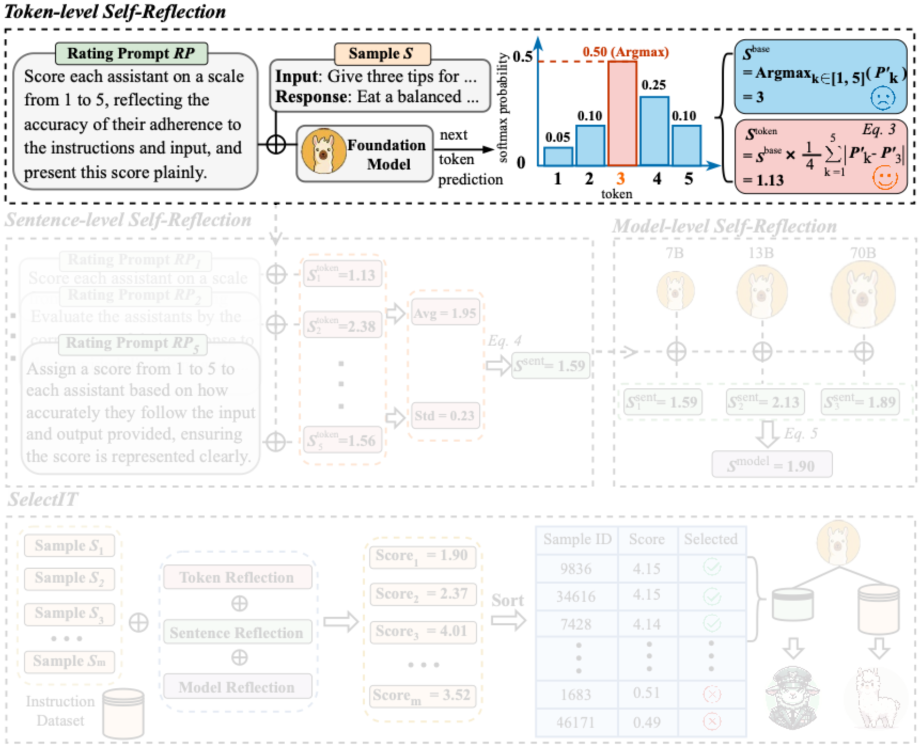
SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection



Overall Framework

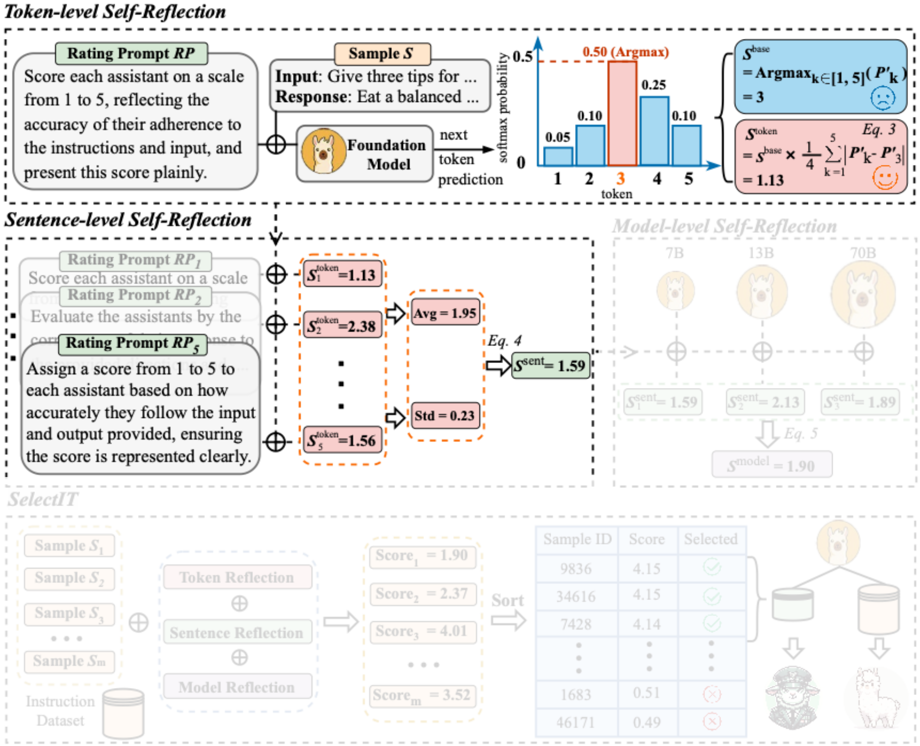
SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection

- SelectIT starts from the LLMs, utilizing their inherent uncertainty to assess the quality of instruction data
  - Token-level: Evaluate the quality using the uncertainty of token probabilities.



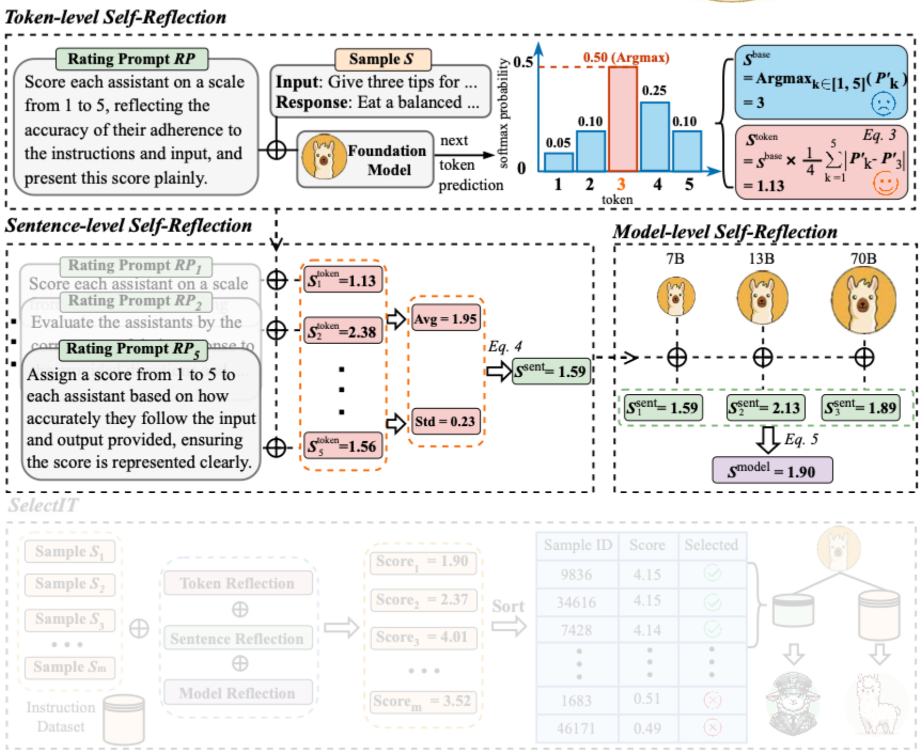
SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection

- SelectIT starts from the LLMs, utilizing their inherent uncertainty to assess the quality of instruction data
  - Token-level: Evaluate the quality using the uncertainty of token probabilities.
  - Sentence-level: Utilize the impact of different prompts on LLMs outputs to improve sample evaluation



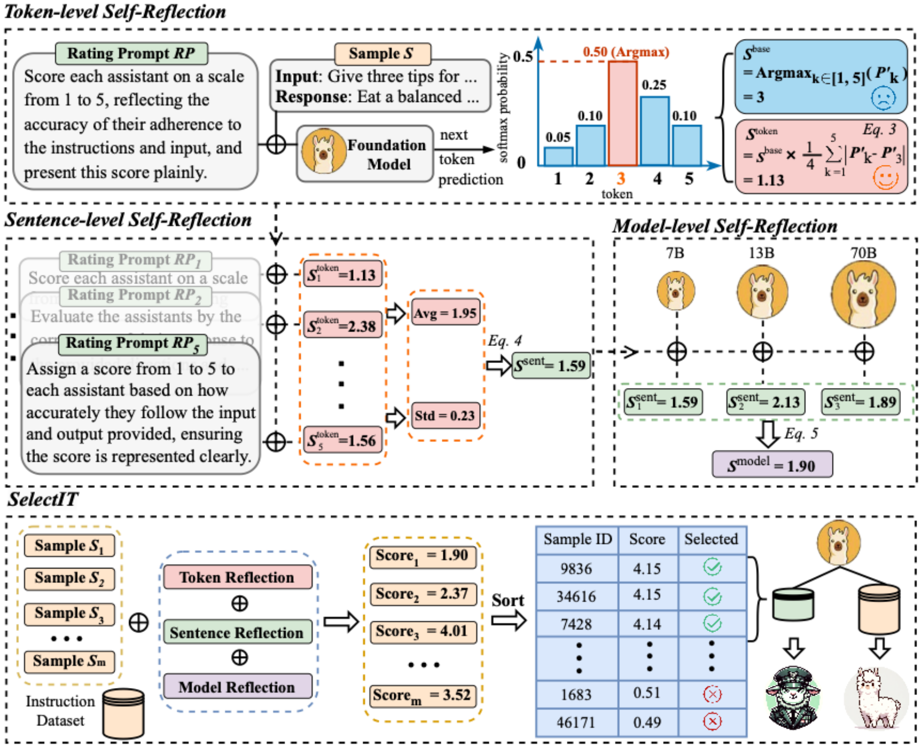
SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection

- SelectIT starts from the LLMs, utilizing their inherent uncertainty to assess the quality of instruction data
  - Token-level: Evaluate the quality using the uncertainty of token probabilities.
  - Sentence-level: Utilize the impact of different prompts on LLMs outputs to improve sample evaluation
  - Model-level: Implement a collaborative decision-making process for data selection based on uncertainty across different LLMs.



SelectIT: Selective Instruction Tuning for LLMs via Uncertainty-Aware Self-Reflection

- SelectIT starts from the LLMs, utilizing their inherent uncertainty to assess the quality of instruction data
  - Token-level: Evaluate the quality using the uncertainty of token probabilities.
  - Sentence-level: Utilize the impact of different prompts on LLMs outputs to improve sample evaluation
  - Model-level: Implement a collaborative decision-making process for data selection based on uncertainty across different LLMs.
  - Selective Alpaca: Apply SelectIT to the instruction dataset of alpaca-gpt4 to propose a new dataset



3

Experiment

Results and Analysis

Datasets and Experiment Settings

- **Tasks and Corresponding Datasets:**
  - Instruction Tuning (IT) and Alpaca-GPT4
- **Evaluation Sets:**
  - MMLU, GSM, BBH, TyDiQA , HumanEval , AlpacaEval
- **Baselines:**
  - Alpaca-GPT4[1]
  - LIMA[2]
  - AlpaGasus[3]
  - From Quantity to Quality[4]

[1] Instruction Tuning with GPT-4. arXiv, 2023.
[2] LIMA: Less Is More for Alignment. NeurIPS, 2023.
[3] AlpaGasus: Training a Better Alpaca with Fewer Data. ICLR, 2024
[4] From Quantity to Quality: Boosting LLM Performance with Self-Guided Data Selection for Instruction Tuning. NAACL, 2024

Results and Analysis

Main Results

- SelectIT can better boost LLaMA-2's performance compared to vanilla IT.
- This enhancement is particularly evident on the BBH and GSM benchmarks.

ID	System	External		MMLU	BBH	GSM	TydiQA	CodeX	AE	Overall	
		Model	Data							AVG	Δ (↑)
Base Model: LLaMA-2-7B				*Implemented Existing Method*							
1	Alpaca-GPT4			46.5	38.4	15.0	43.4	26.8	34.2	34.1	-
2	LIMA	✗	✓	45.4	37.5	14.3	45.1	24.6	33.1	33.3	-0.7
3	1 + AlpaGasus	✓	✗	45.9	39.0	14.5	46.4	27.5	35.4	34.8	+0.7
4	1 + Q2Q	✓	✗	46.9	39.4	15.3	46.7	28.2	35.7	35.4	+1.3
5	1 + Instruction Mining	✓	✓	47.0	39.6	16.5	47.1	28.6	34.4	35.5	+1.5
				Our Proposed Method (Individual)							
6	1 + Token-R	✗	✗	46.8	36.5	14.5	44.6	28.9	35.5	34.5	+0.4
7	1 + Sentence-R	✗	✗	46.9	38.1	16.1	**48.4**	26.9	35.3	35.3	+1.2
8	1 + Model-R	✗	✗	47.3	37.4	16.1	45.3	28.4	**35.8**	35.1	+1.0
				Our Proposed Method (All)							
9	SelectIT (6 + 7 + 8)	✗	✗	**47.4**	**40.6**	**16.8**	47.4	**29.4**	35.7	**36.2**	**+2.2**
Base Model: LLaMA-2-13B				*Implemented Existing Method*							
10	Alpaca-GPT4			**55.7**	46.6	30.5	48.1	40.8	46.5	44.7	-
11	LIMA	✗	✓	54.6	45.3	30.5	51.1	34.1	42.6	43.0	-1.7
12	10 + AlpaGasus	✓	✗	54.1	47.3	31.5	50.6	41.3	46.3	45.2	+0.5
13	10 + Q2Q	✓	✗	55.3	48.5	32.0	50.8	41.3	47.3	45.9	+1.2
14	10 + Instruction Mining	✓	✓	54.1	47.3	32.5	52.6	**43.3**	48.3	46.3	+1.6
				Our Proposed Method (Individual)							
15	10 + Token-R	✗	✗	55.3	47.3	30.5	51.3	39.8	46.2	45.1	+0.4
16	10 + Sentence-R	✗	✗	55.2	48.3	31.0	52.2	42.5	46.3	45.9	+1.2
17	10 + Model-R	✗	✗	55.1	47.5	31.5	52.3	40.2	46.1	45.5	+0.8
				Our Proposed Method (All)							
18	SelectIT (15 + 16 + 17)	✗	✗	**55.7**	**48.9**	**33.0**	**54.1**	42.2	**48.8**	**47.1**	**+2.4**

Table 1: Overall results on IT. "CodeX" and "AE" mean HumanEval and AlpacaEval benchmarks. All the scores are averages of three independent runs with different random seeds.

4

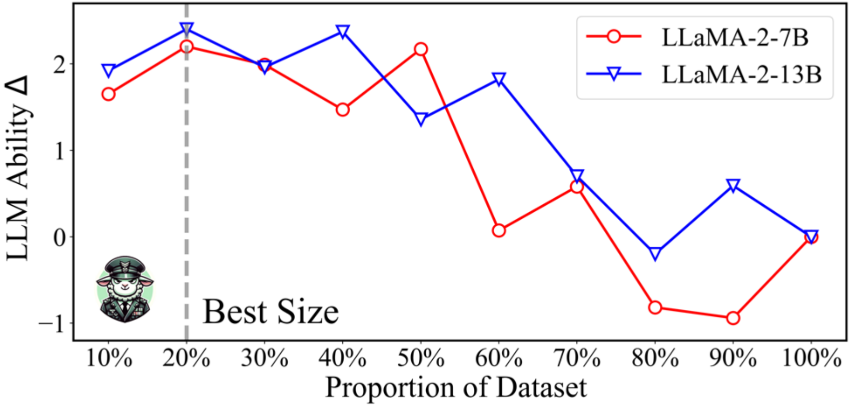
Analysis

Results and Analysis

Ablation Analysis

- **Effect of IT Data Quantity**:

  We opt for 20% for implementing the SelectIT on the Alpaca dataset, base on the tradeoff of training resources, training time, and model performance.



Results and Analysis

Ablation Analysis

Effect of Multiple Rating Prompts

K is a critical parameter for our method, impacting not only the range of scores assigned by the LLMs but also the number of rating prompts. we set K = 5 as the default value in SelectIT.

K	LLaMA-2-7B	LLaMA-2-13B	Overall
3	35.6	46.4	40.5
5	**36.2**	47.1	**41.7**
7	35.7	**47.3**	41.5
9	36.0	46.8	41.4

Table 2: Effect of different K.

Effect of Uncertainty

We ascertain that the α value of 0.2 is optimally suited to establish an effective balance between the sample's quality and the model's uncertainty.

α	MMLU	BBH	GSM	Tydiqa	CodeX	AE	AVG
0.2	47.4	**40.6**	**16.8**	**47.4**	**29.4**	35.7	**36.2**
0.4	**47.9**	39.4	15.5	46.5	**29.4**	**35.8**	35.8
0.6	47.8	39.8	16.5	45.6	29.1	35.1	35.7
0.8	47.6	36.4	16.5	43.6	26.7	35.4	34.4

Table 3: Effect of different α.

Results and Analysis

Ablation Analysis

Effect of Data Imbalance

Base Model	Datasets	Data Size	MMLU	BBH	GSM	Tydiqa	CodeX	AE	Overall	
									AVG	Δ (↑)
LLaMA-2-7B	LIMA	1K	45.4	37.5	14.3	45.1	24.6	33.1	33.3	-
	Selective Alpaca	1K	**46.6**	**41.3**	**14.5**	**46.2**	**30.6**	**33.8**	**35.5**	+2.2
	AlpaGasus	9K	45.9	39.0	14.5	**46.4**	27.5	35.4	34.8	-
	Selective Alpaca	9K	**47.2**	**41.3**	**18.5**	47.6	**28.3**	**35.4**	**36.4**	+1.6

Table 5: Results on IT for different datasets with the same number of instances.

- When facing the same amount of data, SelectIT can still demonstrate better performances, which further illustrates its effectiveness.

Results and Analysis

Robustness across Models, Datasets and Domains

Various Instruction Tuning Datasets

Datasets	Data Size	MMLU	BBH	GSM	Tydiqa	CodeX	AE	Overall	
								AVG	Δ (↑)
WizardLM	143K	43.8	37.8	10.0	41.2	25.2	**35.3**	32.2	-
WizardLM + SelectIT	28.6K	**45.1**	**40.1**	**11.0**	**43.1**	**27.5**	34.7	**33.6**	+1.4
Orca-GPT4	1M	40.1	35.6	13.0	**46.0**	23.3	**38.1**	32.7	-
Orca-GPT4 + SelectIT	0.2M	**43.9**	**38.7**	**16.5**	42.0	**27.7**	37.4	**34.4**	+1.7

Table 7: Results of IT with various IT datasets.

- SelectIT consistently enhances the performance of the model on both the WizardLM and Orca-GPT4 datasets.

Results and Analysis

Robustness across Models, Datasets and Domains

Various Domain-specific Tasks

- SelectIT is a versatile and scalable method, effective not only for IT data selection but also for domain-specific tasks like MT.

Method	Size	ALL	
		COMET	BLEU
SoTA Models			
NLLB (Costa-jussà et al., 2022)	54B	78.8	26.3
GPT-3.5	-	85.6	34.8
GPT-4	-	85.8	35.1
Existing Method			
LLaMA-2 (Touvron et al., 2023b)	7B	76.5	21.1
TIM (Zeng et al., 2023)	7B	79.1	26.4
SWIE (Chen et al., 2023b)	7B	80.6	27.6
BigTranslate (Yang et al., 2023)	13B	78.8	21.9
Bayling (Zhang et al., 2023)	13B	82.0	27.8
Our Implemented Method			
ALMA	7B	83.2	29.7
w/ SelectIT	7B	**83.7**	**30.5**
ALMA	13B	83.7	31.5
w/ SelectIT	13B	**84.2**	**32.2**

Table 8: The overall results on MT LLMs.

Efficiency of SelectIT

- Although Selective Alpaca is selected by the LLaMA-2 models, it is also applicable to the Mistral-7B and LLaMA-3-8B.

Method	Speed	Time	Cost
ChatGPT API	0.76 it/s	19.07h	$52.02
GPT4 API	0.37 it/s	38.98h	$2871.56
SelectIT	**9.34 it/s**	**5.80h**	**$26.68**

Table 9: Comparison of selection efficiency.

Results and Analysis

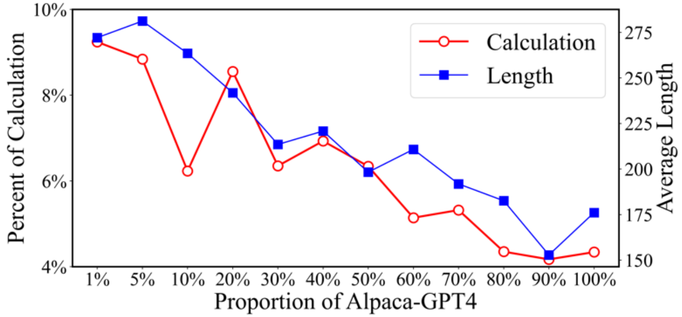
Insights of Selective Data

- **Different Selection Strategies**
  - SelectIT can significantly better improve the abilities of LLMs than random selection methods.

Method	LLaMA-2		ALMA		Δ (↑)
	7B	13B	7B	13B	
Full Dataset	34.1	44.2	29.7	31.5	-
w/ Random (Full)	34.1	45.1	29.3	31.0	0.0
w/ Random (Unselected)	34.6	44.3	29.1	31.2	-0.4
w/ Length	35.5	47.1	30.1	31.8	+5.0
w/ SelectIT	**36.2**	**47.1**	**30.5**	**32.2**	**+6.5**

Table 10: Comparasion with variants.

- **Insights of High-Quality Data in SelectIT**
  - SelectIT can reasonably rank samples based on their characteristics.

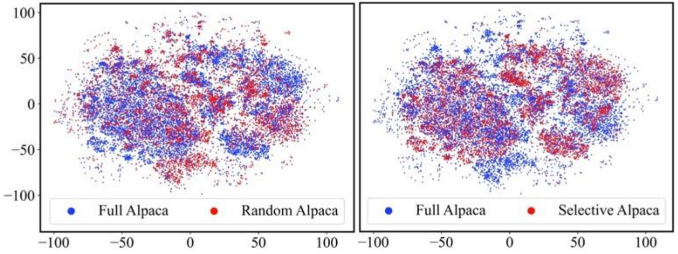


Results and Analysis

Insights of Selective Data

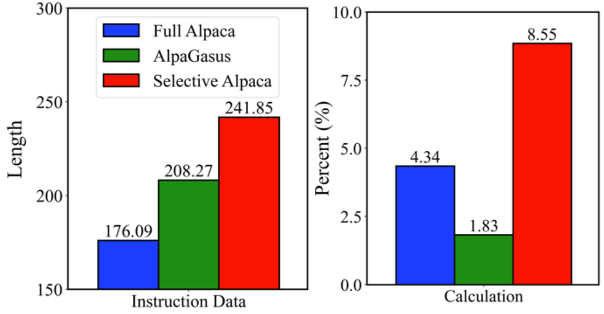
Data Representation Analysis

- Selective Alpaca data are mostly concentrated around the center, indicating that our dataset predominantly contains high-quality data near the center



Data Characteristic Analysis

- SelectIT tends to select high-quality mathematical data, providing a solid explanation for the observed improvement in the reasoning abilities of LLMs.



5

Conclusion

Conclusion

- We propose SelectIT, a novel IT data selection method which exploits the uncertainty of LLMs without using additional resources.

- We introduce a curated IT dataset, Selective Alpaca, by selecting the high-quality IT data from the Alpaca-GPT4 dataset.

- Our analysis suggests that longer and more computationally intensive IT data may be more effective, offering a new perspective on the characteristics of optimal IT data.

Thanks for your listening!