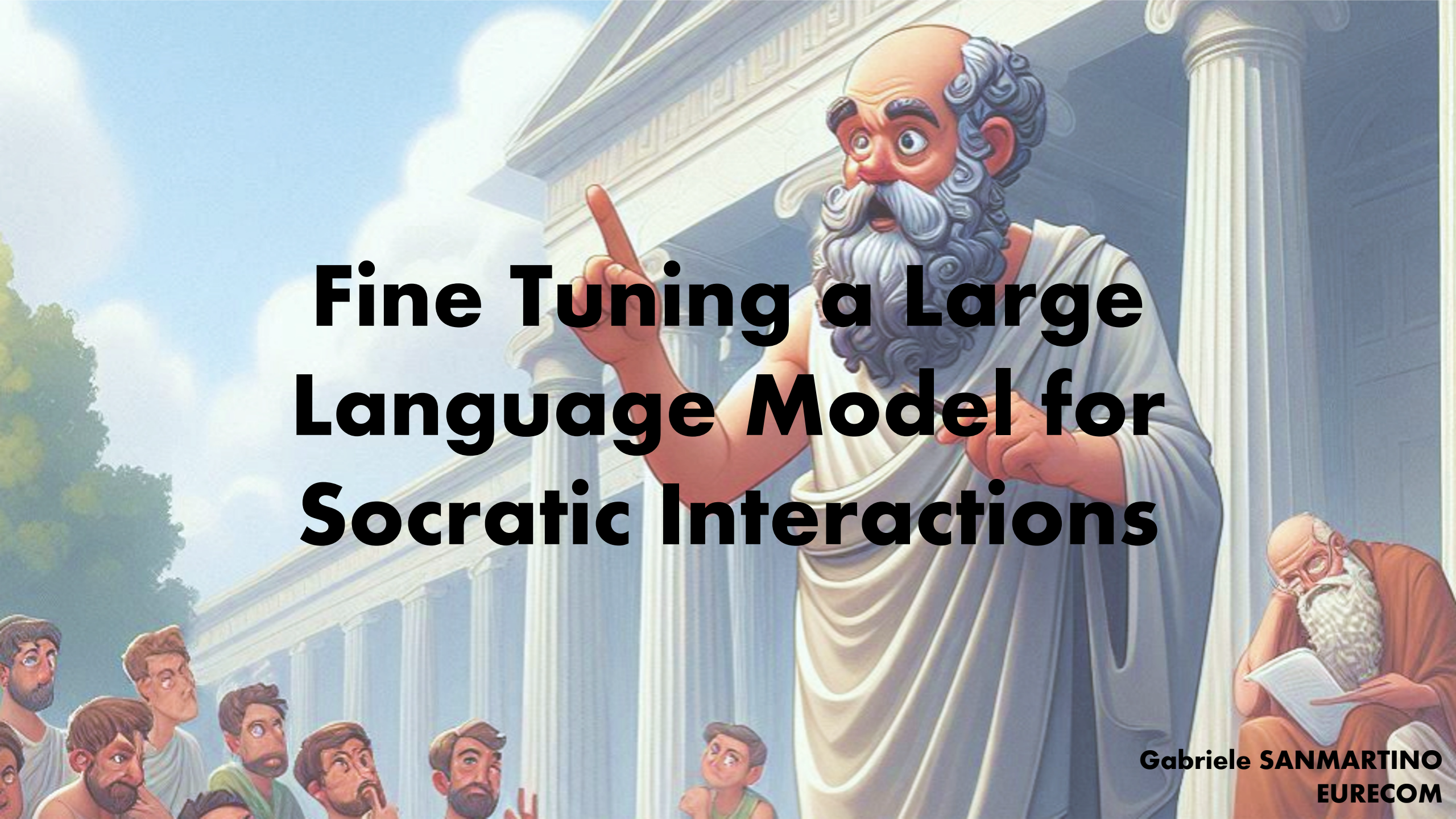
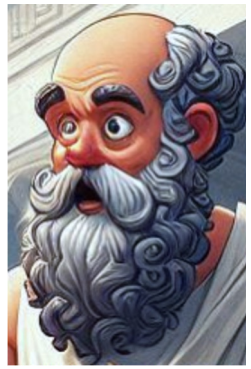


Fine Tuning a Large Language Model for Socratic Interactions

Gabriele SANMARTINO
EURECOM

Professor, how can I solve $x^2 - 1 = 0$?

GPT 4 Turbo



To solve the equation $x^2 - 1 = 0$, you can follow these steps:

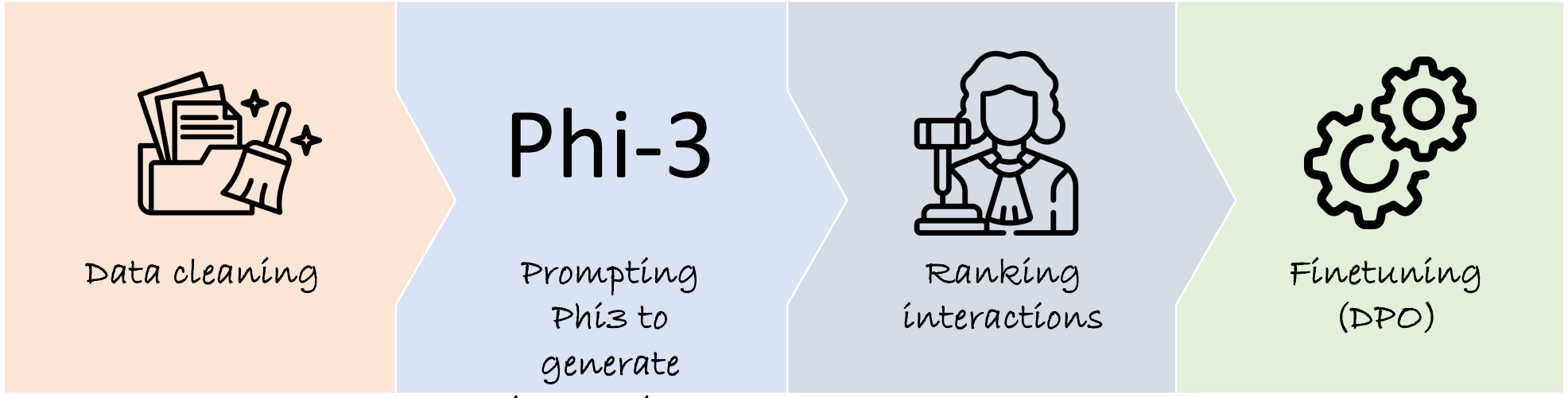
[...]

So, the solutions to the equation $x^2 - 1 = 0$ are:

$x = 1$ and $x = -1$.

How does this equation resemble any identities or formulas you have studied before, such as $(a + b)(a - b) = a^2 - b^2$?

Training pipeline, in a Nutshell



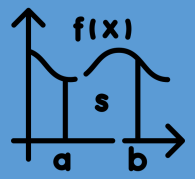
Datasets

TutorChat

Science textbooks

GPT-3.5-turbo/GPT-4-turbo (teacher) vs. GPT-3.5-turbo/GPT-4-turbo (student)

MathDial



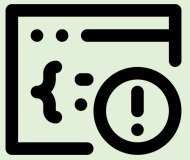
Problem statement

Student's answer

Correct answer

Human (teacher) vs. GPT-3.5-turbo (student)

Debugging



Problem statement

Student's code

Bug descr.

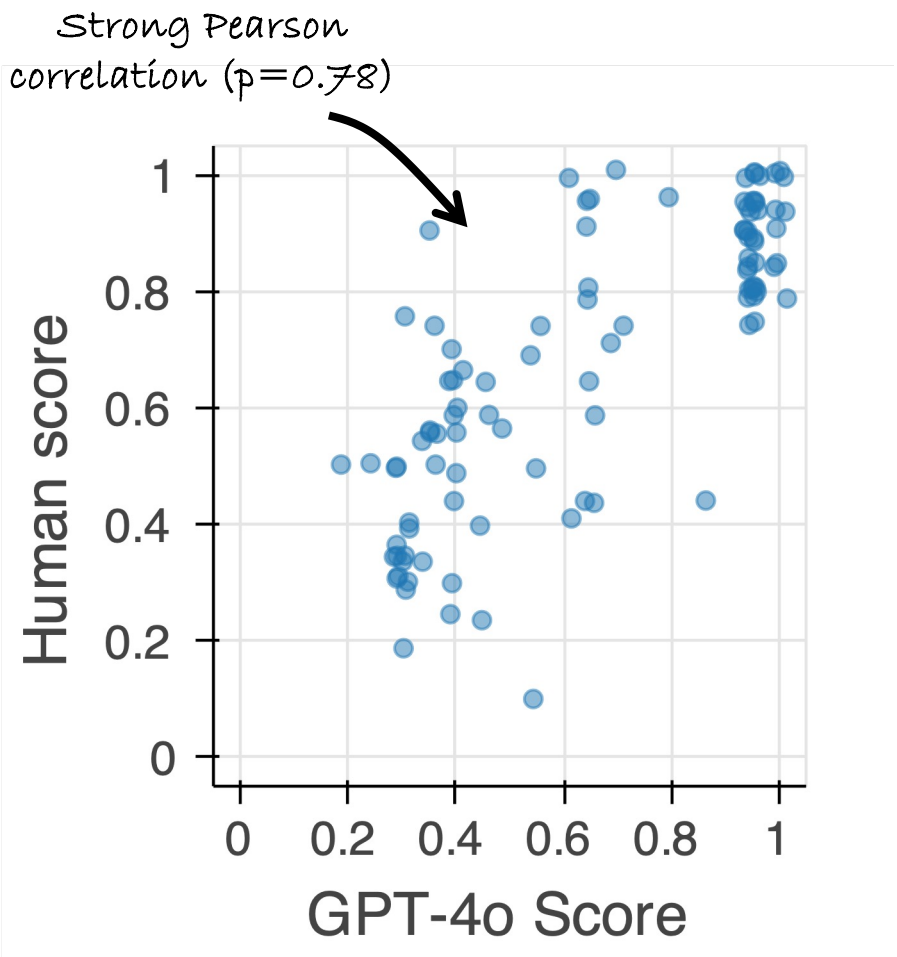
Human (teacher) vs. Human (student)

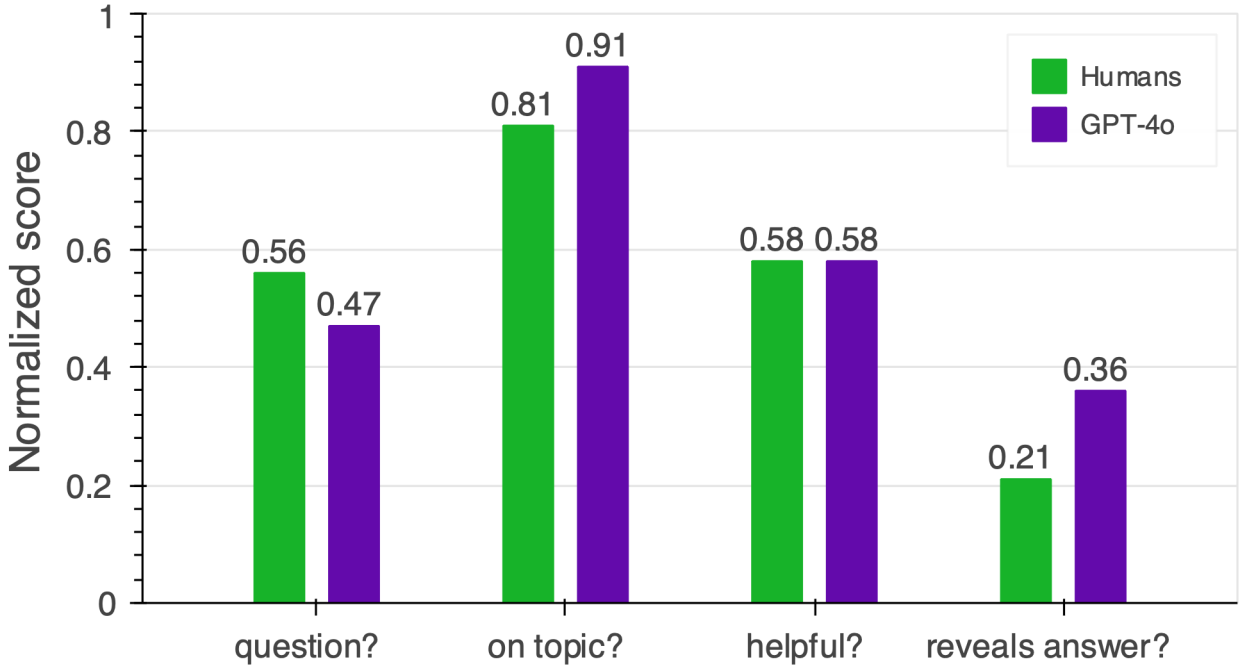
Evaluation criteria

Does the answer contain a **question**?	Does the answer **reveal** the correct solution?	How much **on topic** the answer is with respect to the conversation?	How **helpful** is the answer to reach the solution?
Boolean	Boolean	Scale 1 to 5	Scale 1 to 5

Summary score is the average of these four criteria

Validating assessment automation

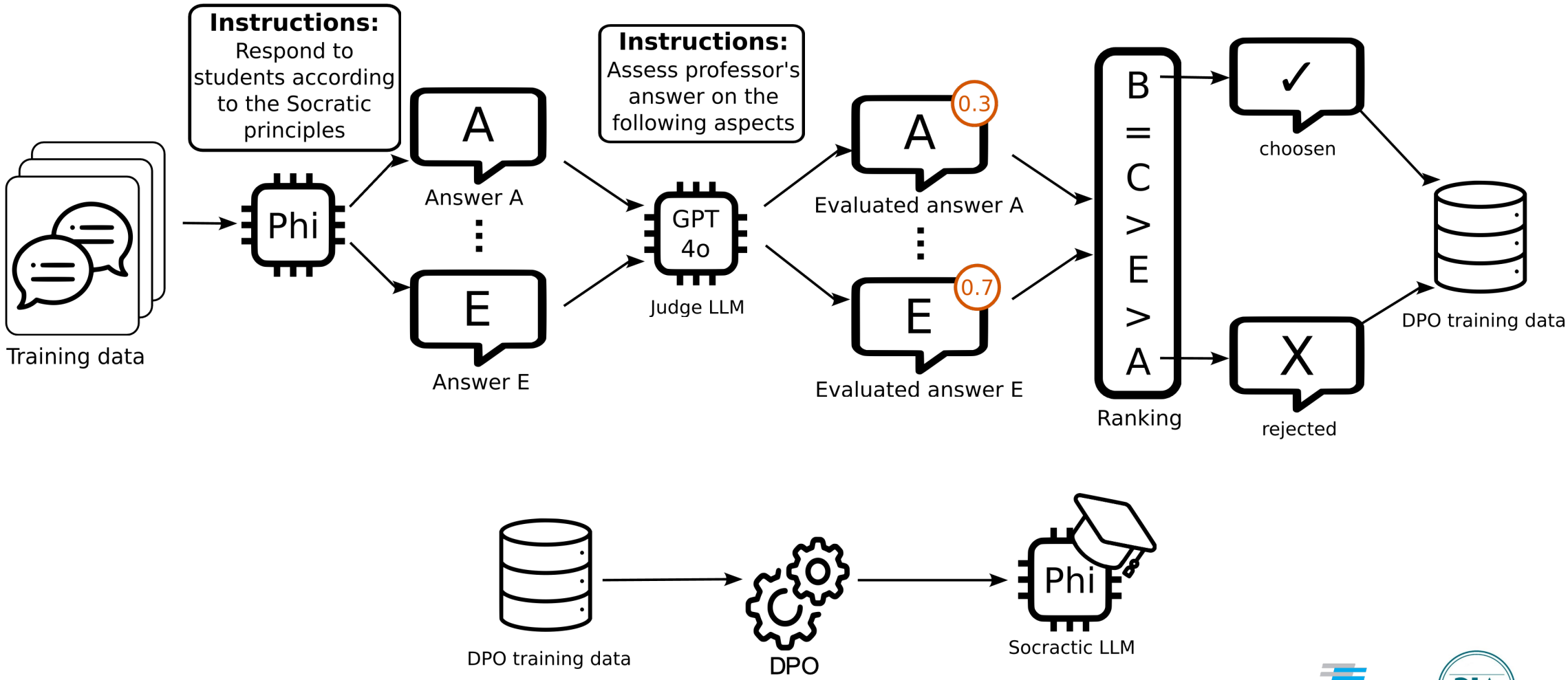




Optimization with DPO

- Stable
- No need to train a reward model
- No need to load a reward model into GPUs during fine tuning

Training pipeline



Results

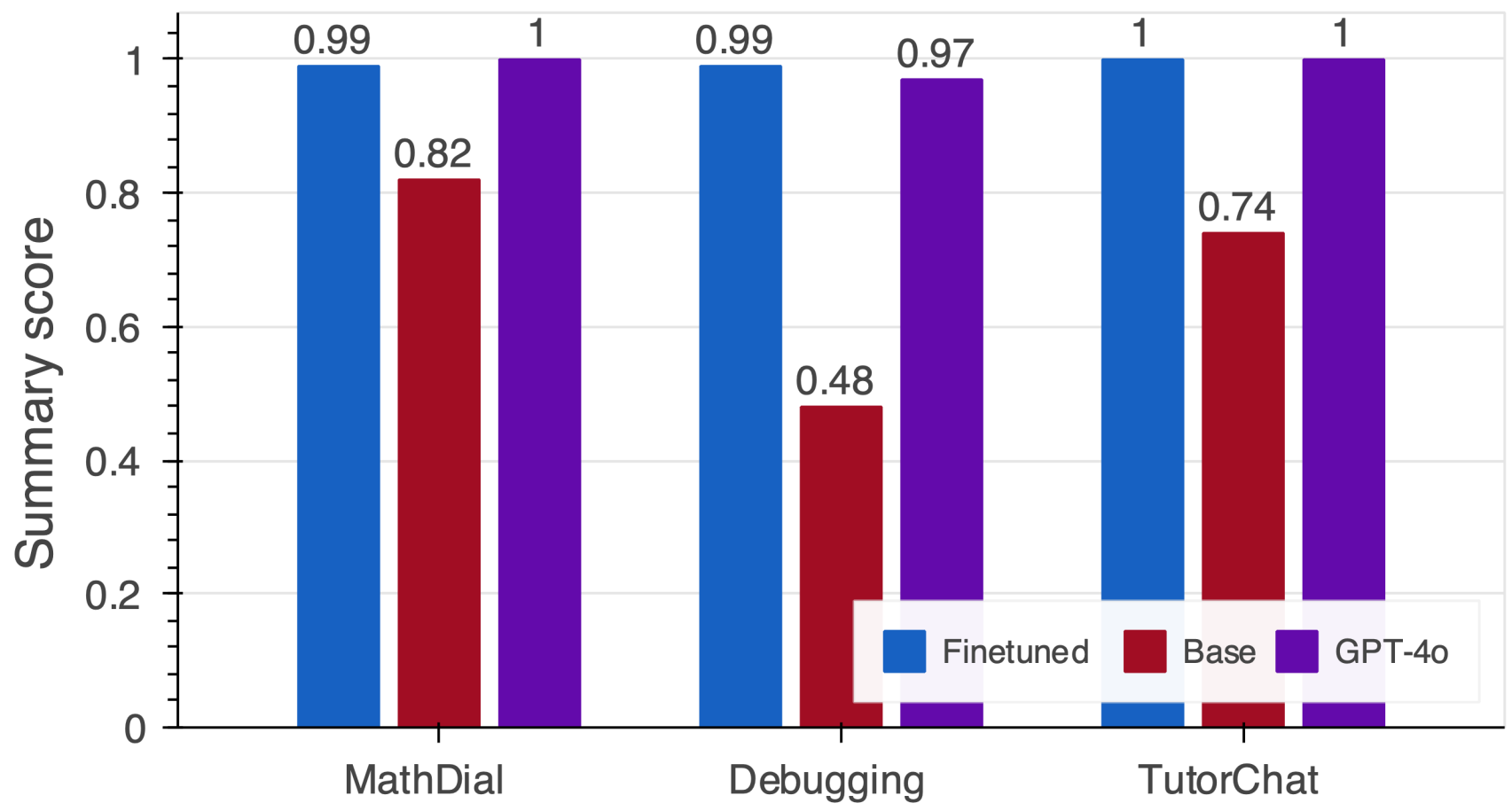
Our best model

Test Dataset	Train Dataset		
	mathdial	tutorchat	debugging
mathdial	0.98	0.99	0.98
tutorchat	0.98	1.00	0.96
debugging	0.97	0.99	0.94

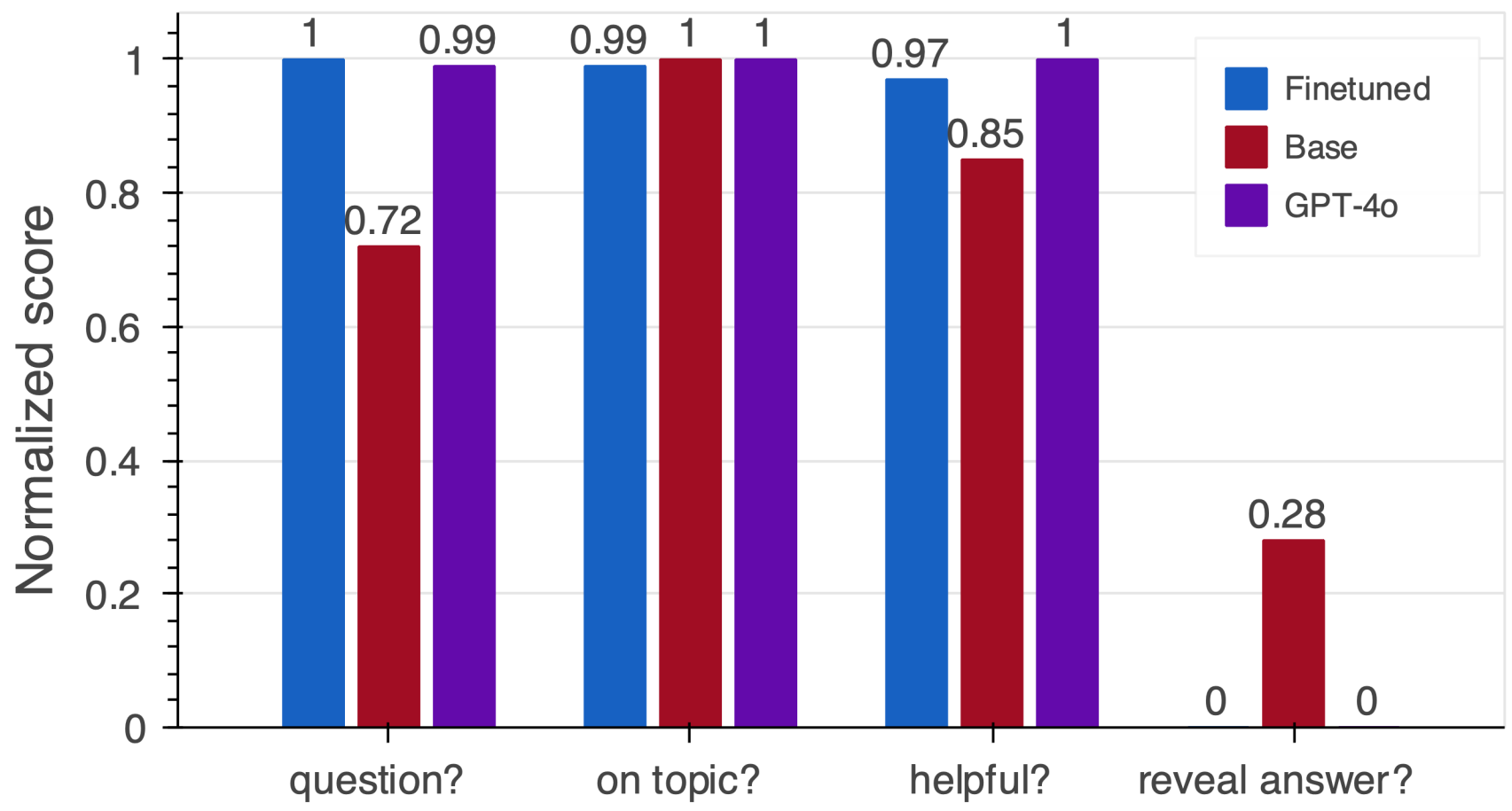
?

Table 1: Summary scores on 100 samples for various training and testing datasets.

Results


Summary score
1
0.8
0.6
0.4
0.2
0
0.99
0.82
1
0.99
0.48
0.97
1
0.74
1
MathDial
Debugging
TutorChat
Finetuned
Base
GPT-4o

Results


Normalized score
1
0.8
0.6
0.4
0.2
0
1
0.72
0.99
0.99
1
1
0.97
0.85
1
0
0.28
0
question?
on topic?
helpful?
reveal answer?
Finetuned
Base
GPT-4o

Contributions

- Making available Socratic models for the community
- Introduced a rule-based assessment technique for the Socratic behavior
- State-of-the-art performance with a cheap model
- Blog: https://giovannigatti.github.io/socratic-llm/


https://github.com/GiovanniGatti/socratic-llm


https://www.youtube.com/watch?v=Phh4PhZOIE0


https://huggingface.co/eurecom-ds/Phi-3-mini-4k-socratic


https://ollama.com/eurecom-ds/phi-3-mini-4k-socratic

Thanks to the team

Gabriele Sanmatino

Prof. Raphaël Troncy

Giovanni Gatti Pinheiro

Giulia Bonino

Prof. Paolo Papotti

Prof. Pietro Michiardi