



Improving Multi-Domain Learning by Balancing Batches With Domain Information

Alexandre Thurow Bender

Advisor: Prof. Dr. Ricardo Matsumura Araujo
Co-Advisor: Prof. Dr. Ulisses Brisolara Corrêa

Introduction

- Limited Data Quantity
- Representative Datasets
- Data Acquisition

Domains



Domains



Domains



Domains



Domains



Domain B

Domain A

Datasets

Amazon



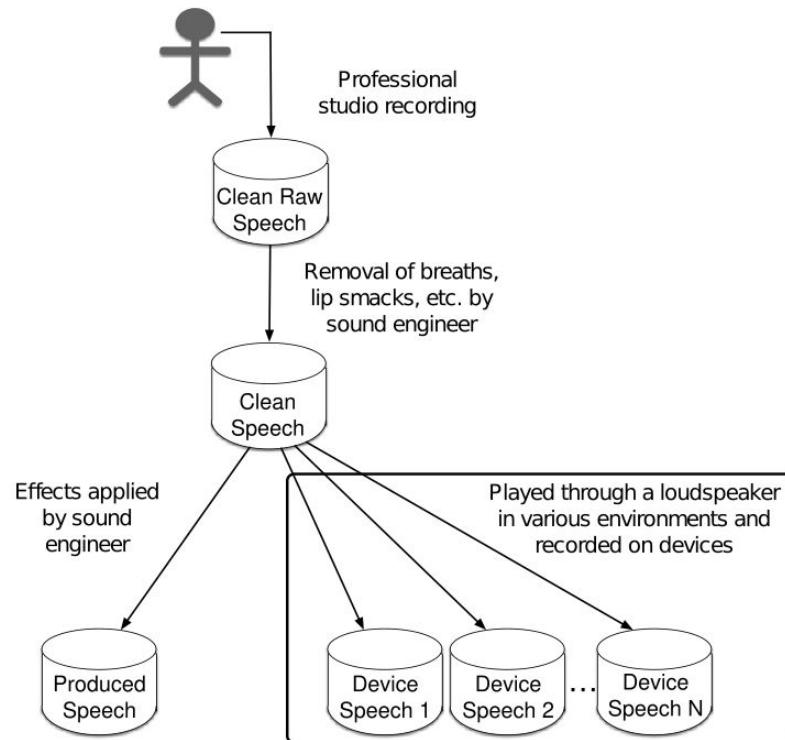
DSLR



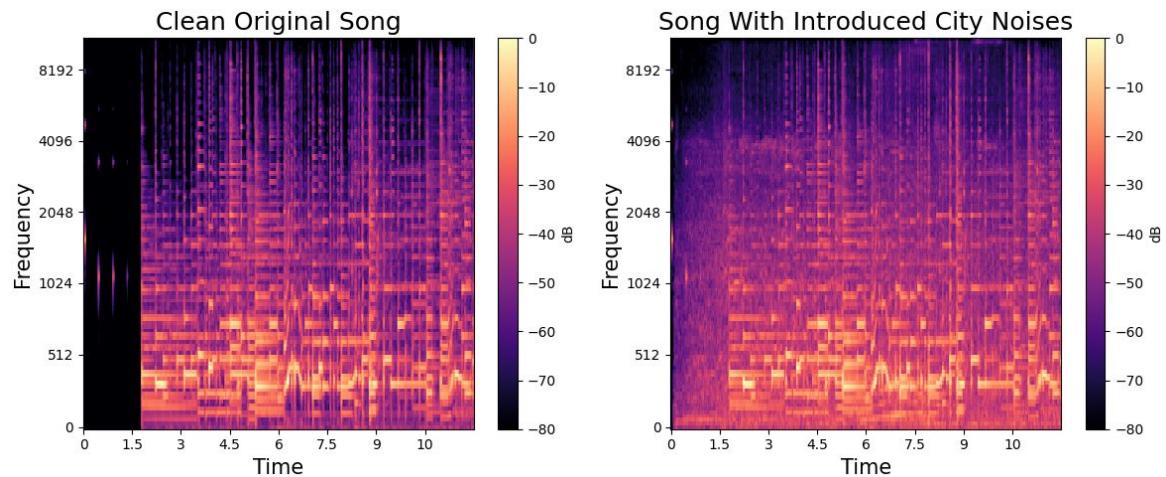
Webcam



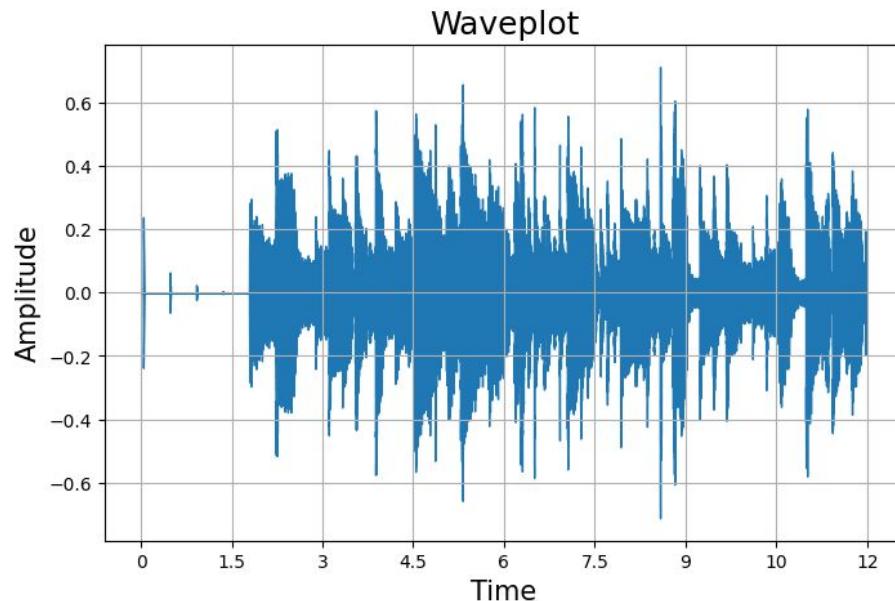
Datasets



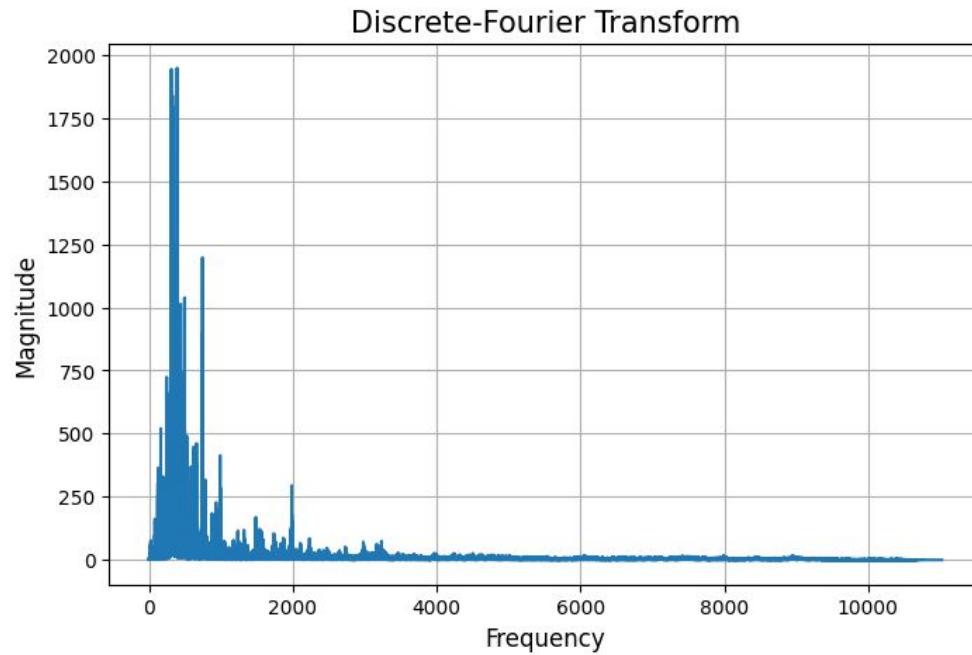
Differentiating Domains



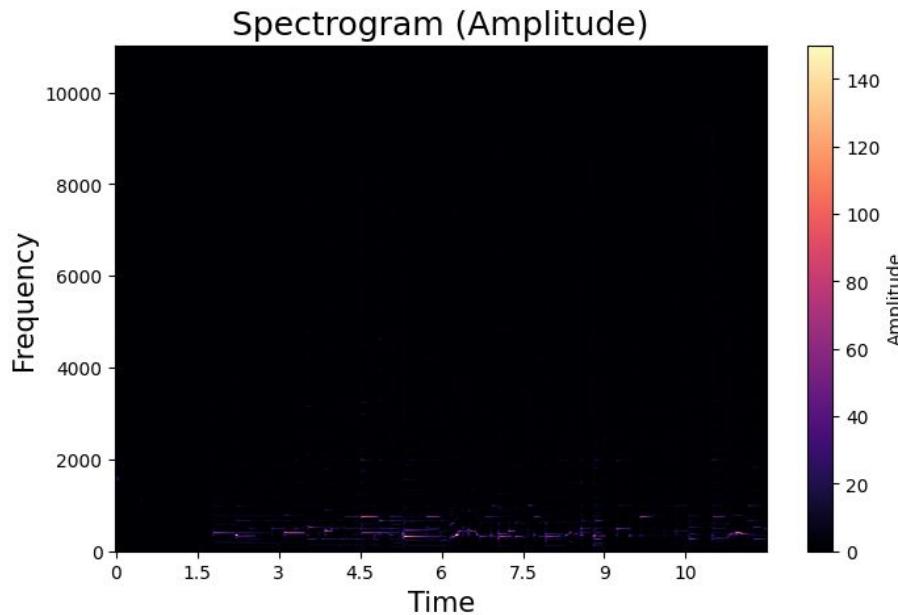
Audio Processing



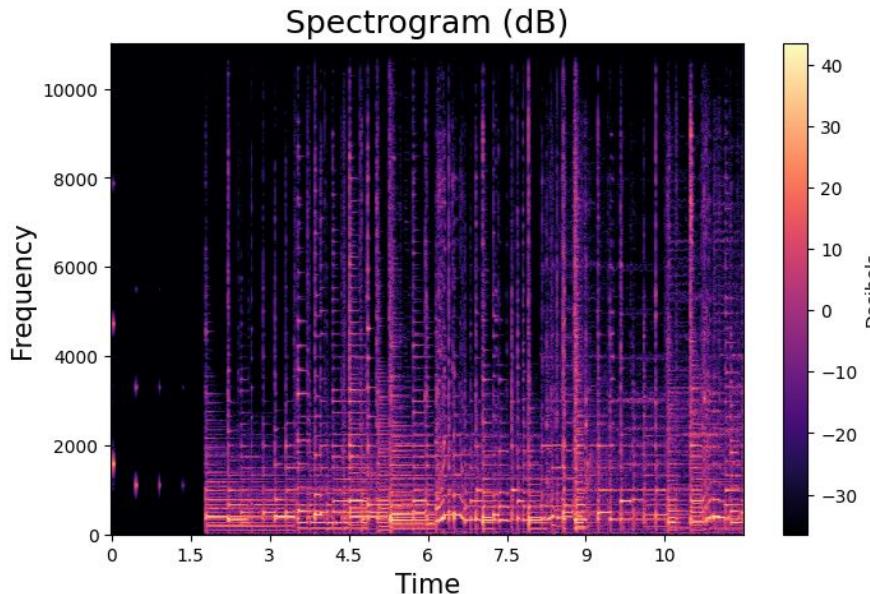
Audio Processing



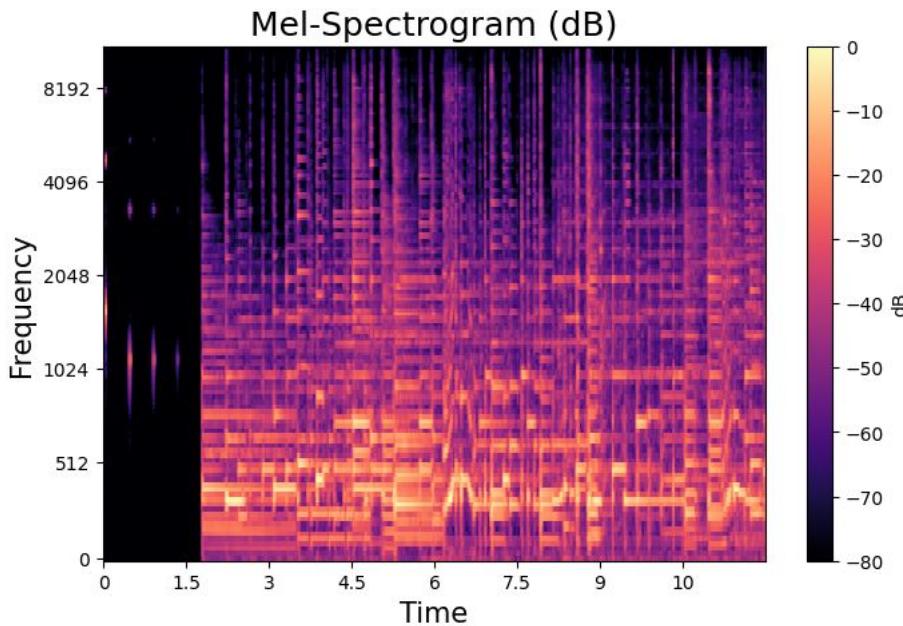
Audio Processing



Audio Processing

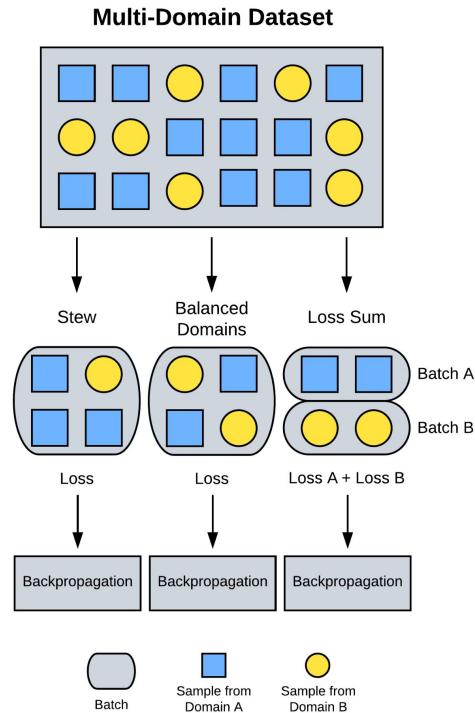


Audio Processing

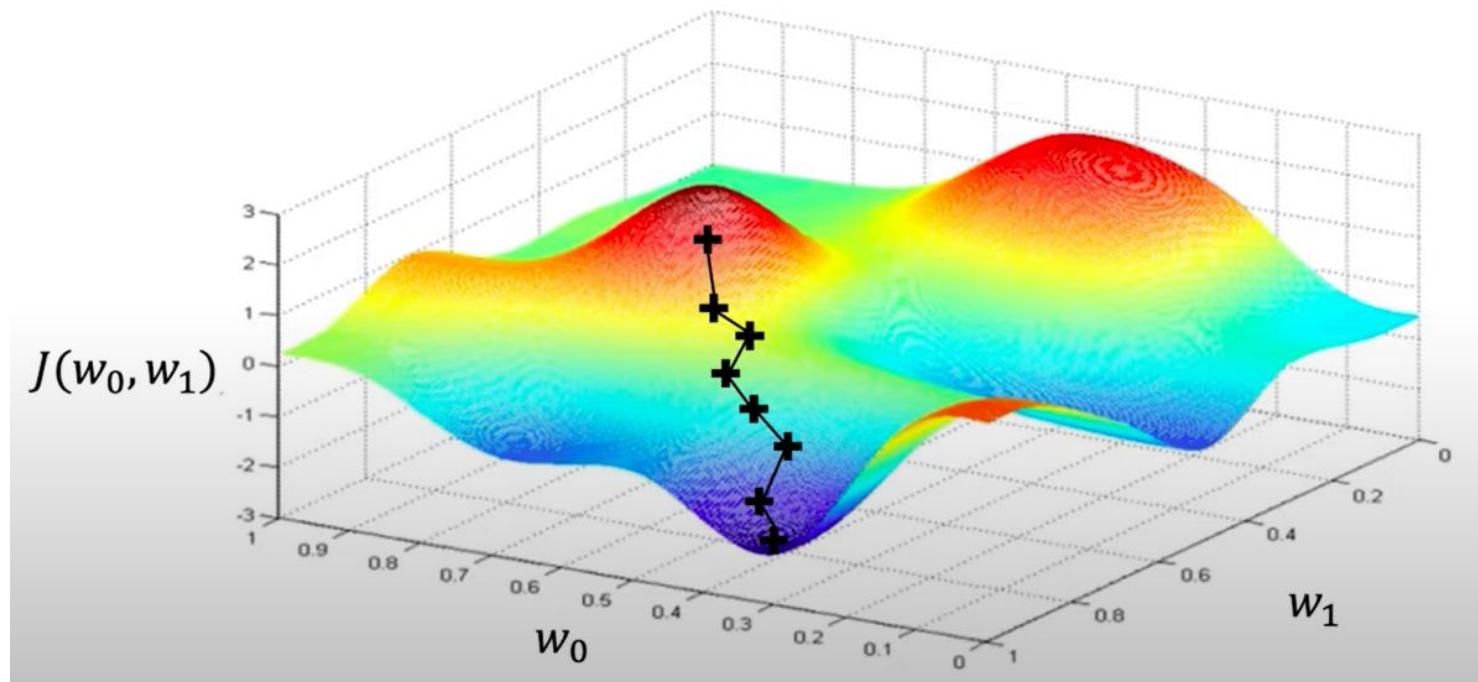


$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right)$$

Methods



Loss Functions



Results – Audio Classification

Domain	Stew	Balanced Domains	Loss_Sum
CLEAN	0.923 ± 0.011	0.908 ± 0.022	0.952 ± 0.010
IPAD_BALCONY1	0.854 ± 0.014	0.853 ± 0.023	0.888 ± 0.013
IPAD_BEDROOM1	0.919 ± 0.013	0.912 ± 0.013	0.944 ± 0.008
IPAD_CONFROOM1	0.931 ± 0.010	0.923 ± 0.015	0.952 ± 0.011
IPAD_CONFROOM2	0.931 ± 0.012	0.929 ± 0.012	0.955 ± 0.007
IPADFLAT_CONFROOM1	0.901 ± 0.012	0.906 ± 0.018	0.930 ± 0.009
IPADFLAT_OFFICE1	0.909 ± 0.011	0.901 ± 0.016	0.938 ± 0.013
IPAD_LIVINGROOM1	0.919 ± 0.008	0.920 ± 0.018	0.952 ± 0.008
IPAD_OFFICE1	0.897 ± 0.019	0.891 ± 0.019	0.923 ± 0.014
IPAD_OFFICE2	0.883 ± 0.013	0.883 ± 0.020	0.919 ± 0.009
IPHONE_BALCONY1	0.825 ± 0.022	0.811 ± 0.025	0.853 ± 0.016
IPHONE_BEDROOM1	0.873 ± 0.012	0.863 ± 0.015	0.913 ± 0.013
IPHONE_LIVINGROOM1	0.889 ± 0.011	0.877 ± 0.023	0.913 ± 0.011
PRODUCED	0.901 ± 0.014	0.887 ± 0.024	0.941 ± 0.011
Average F1-Score	0.897 ± 0.010	0.890 ± 0.016	0.927 ± 0.008

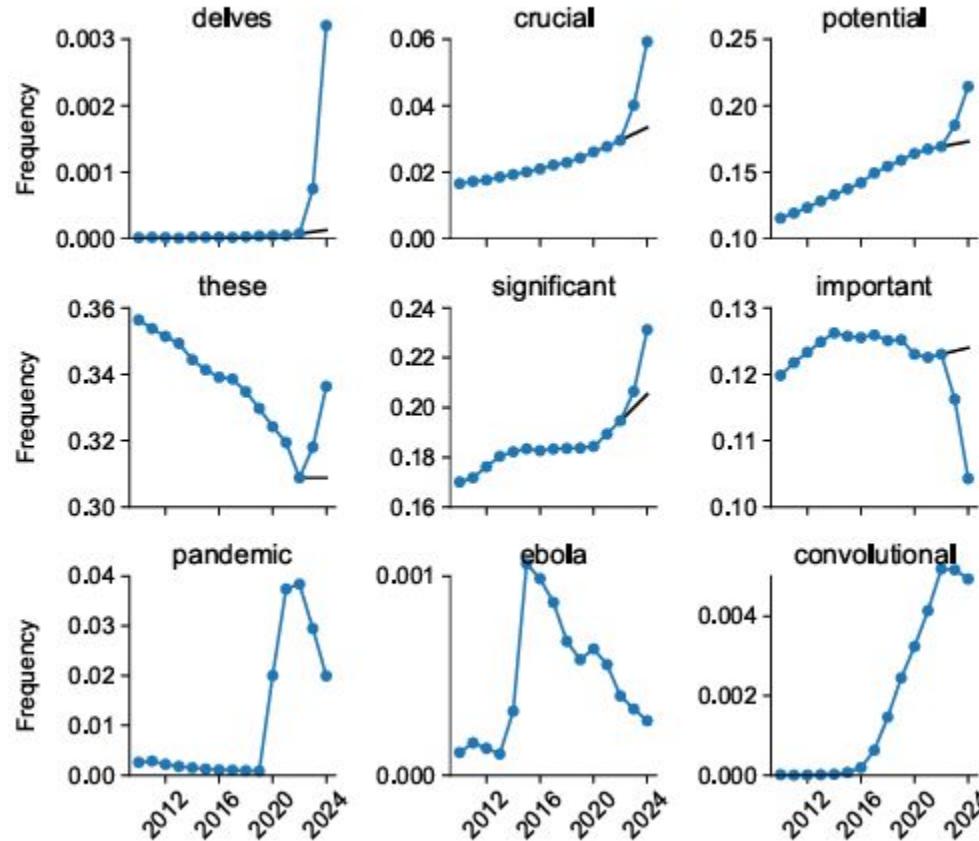


The TARDIS Operational Manual: Creating and Evaluating a Synthetic Dataset for Retrieval-Augmented Generation

Alexandre Thurow Bender

Advisor: Prof. Dr. Ricardo Matsumura Araujo
Co-Advisor: Prof. Dr. Ulisses Brisolara Corrêa

Why LLMs?



Oh no!

How many times was the UFPel Campus Anglo transported across the Canal São Gonçalo in early 2004?



The UFPel Campus Anglo was transported across the Canal São Gonçalo **three times** in early 2004. This was done using a barge, as part of a significant operation to move the historical Anglo building to its current location. The entire process required meticulous planning and engineering expertise to ensure the safety of the structure during the crossings.



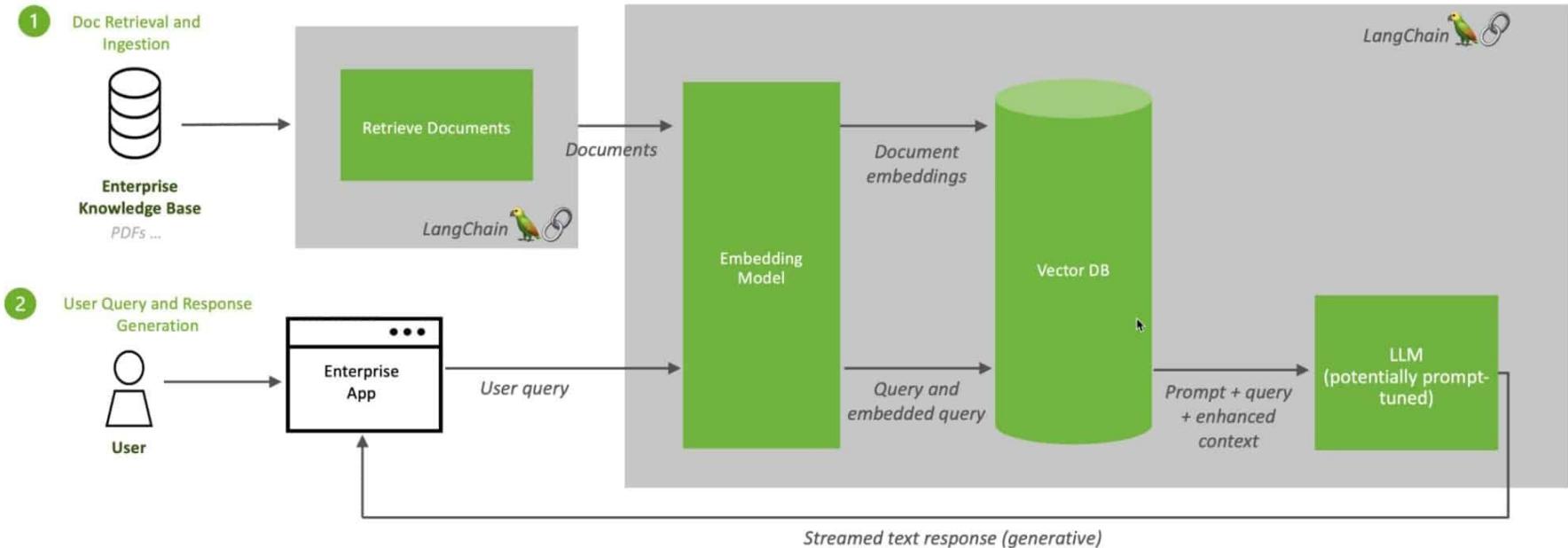
Message ChatGPT



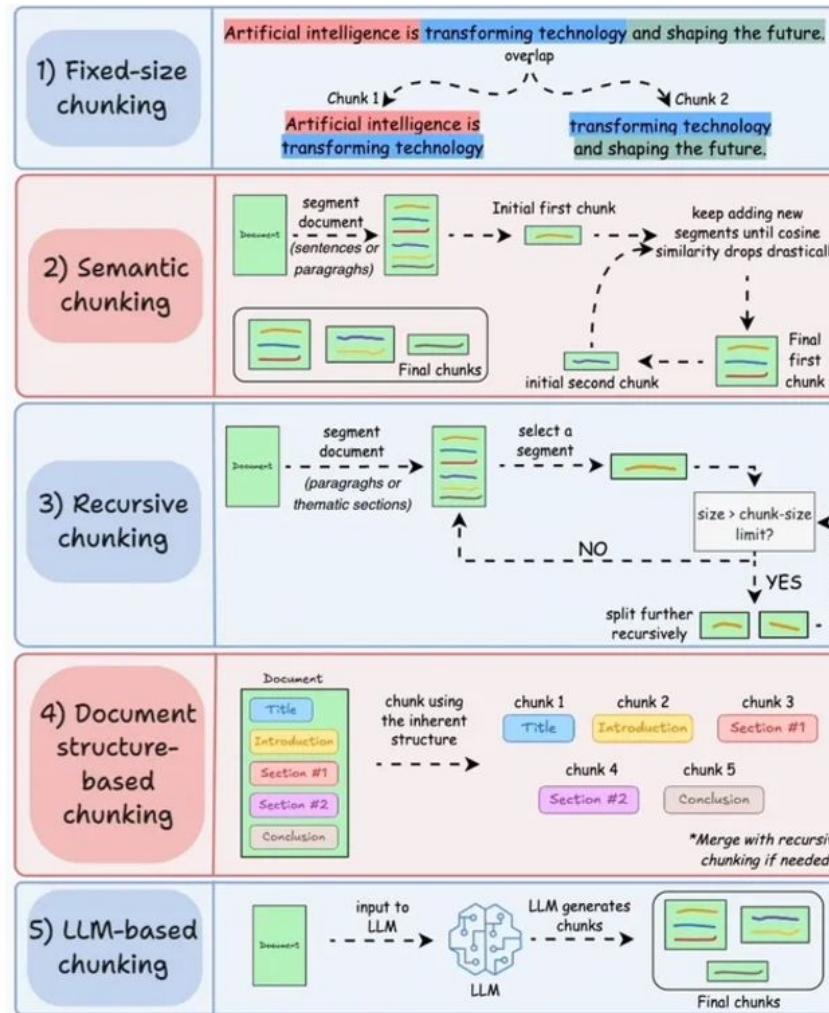
ChatGPT can make mistakes. Check important info.

RAG!

Retrieval Augmented Generation (RAG) Sequence Diagram



RAG... :(



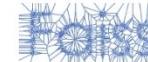
RAG... :(

Choosing a Vector Database

Vector Databases



Vector Libraries



Vector-Capable NoSQL Databases



Vector-Capable SQL Databases



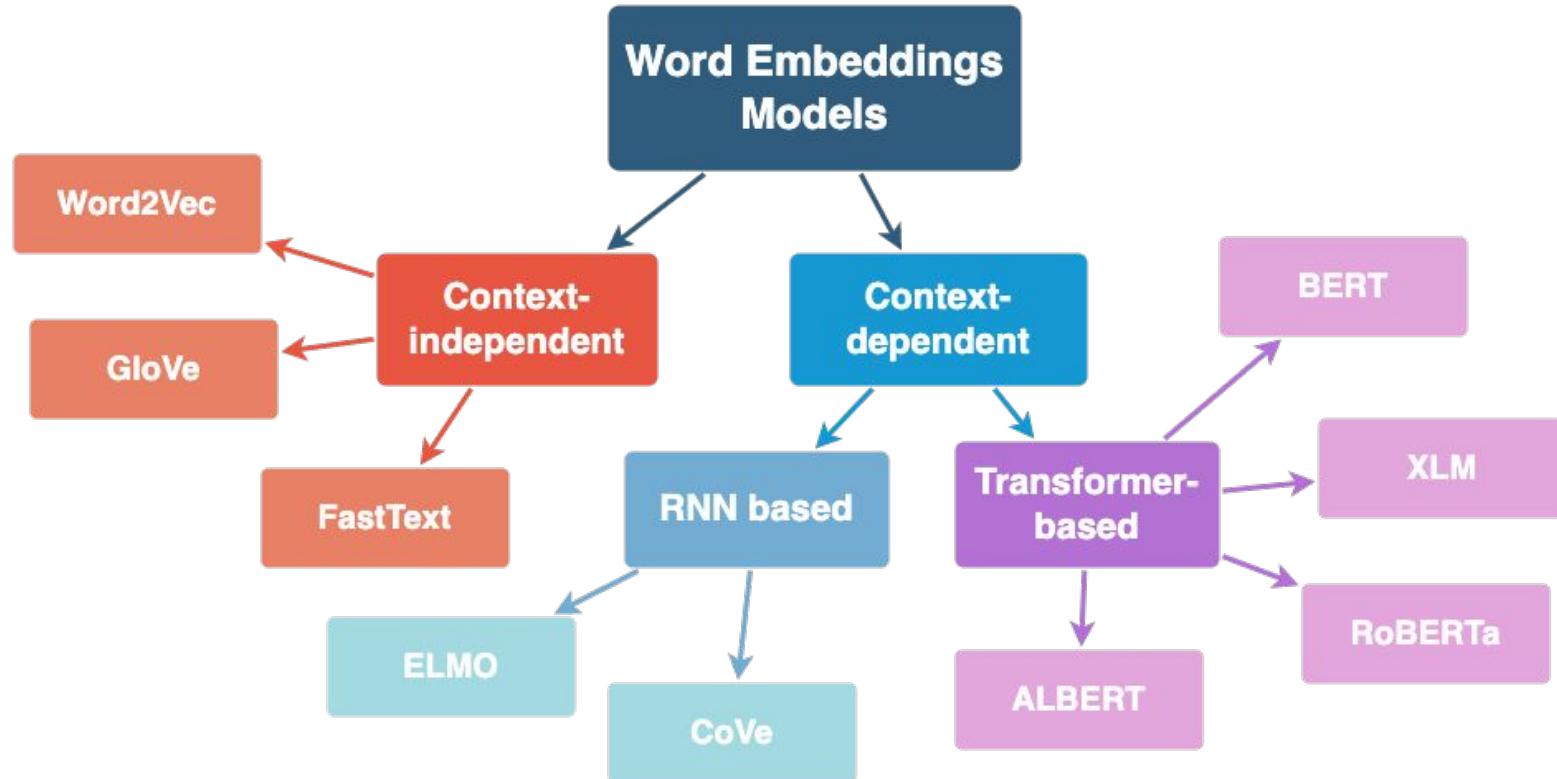
Timescale



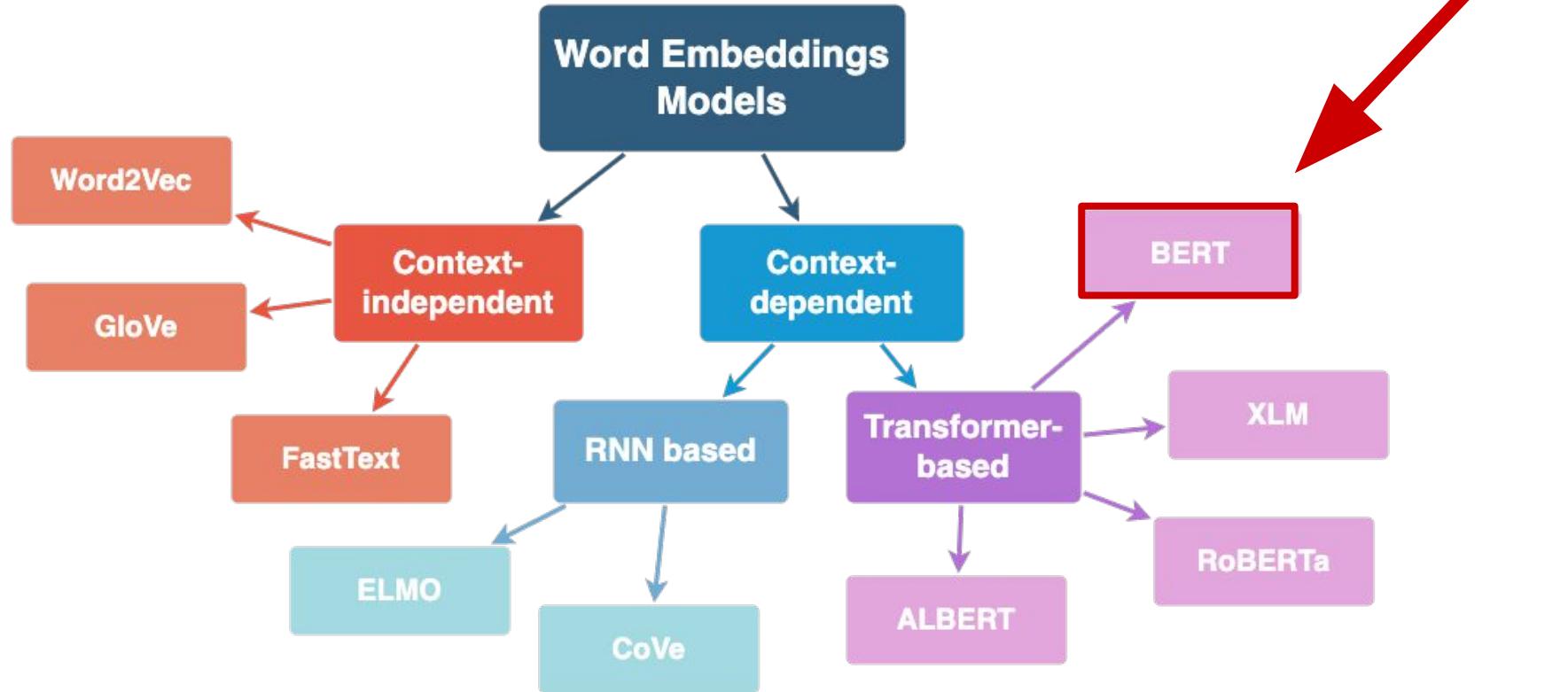
Text Search Databases



RAG... :(



RAG... :(



RAG... :(



BERT community

Follow 145

AI & ML interests

This organization is maintained by the transformers team at Hugging Face and contains the historical (pre-"Hub") BERT checkpoints.

Team members 3



Models 15



↑ Sort: Recently updated

google-bert/bert-large-cased-whole-word-masking

Fill-Mask • Updated Apr 10 • ↓ 6.09k • ❤ 15

google-bert/bert-large-uncased-whole-word-masking-f...

Question Answering • Updated Feb 19 • ↓ 115k • ⚡ • ❤ 172

google-bert/bert-large-uncased-whole-word-masking

Fill-Mask • Updated Feb 19 • ↓ 28.5k • ❤ 19

google-bert/bert-large-uncased

Fill-Mask • Updated Feb 19 • ↓ 1.82M • ❤ 116

google-bert/bert-large-cased-whole-word-masking-fin...

Question Answering • Updated Feb 19 • ↓ 180k • ❤ 1

google-bert/bert-large-cased

Fill-Mask • Updated Feb 19 • ↓ 1.45M • ❤ 31

google-bert/bert-base-uncased

Fill-Mask • Updated Feb 19 • ↓ 69.9M • ❤ 1.92k

google-bert/bert-base-multilingual-uncased

Fill-Mask • Updated Feb 19 • ↓ 2.43M • ❤ 108

google-bert/bert-base-multilingual-cased

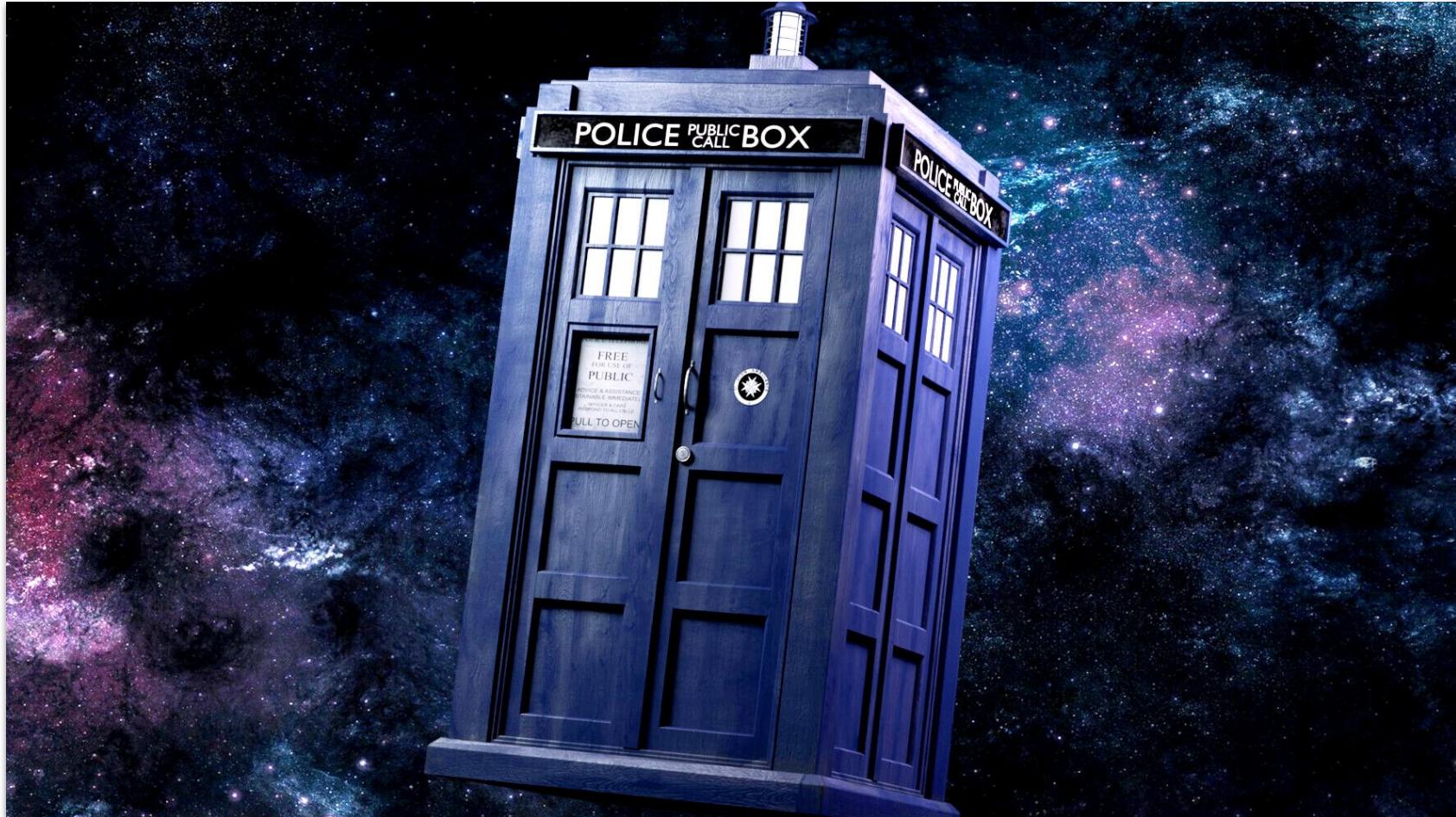
Fill-Mask • Updated Feb 19 • ↓ 5.35M • ❤ 449

google-bert/bert-base-german-dbmdz-uncased

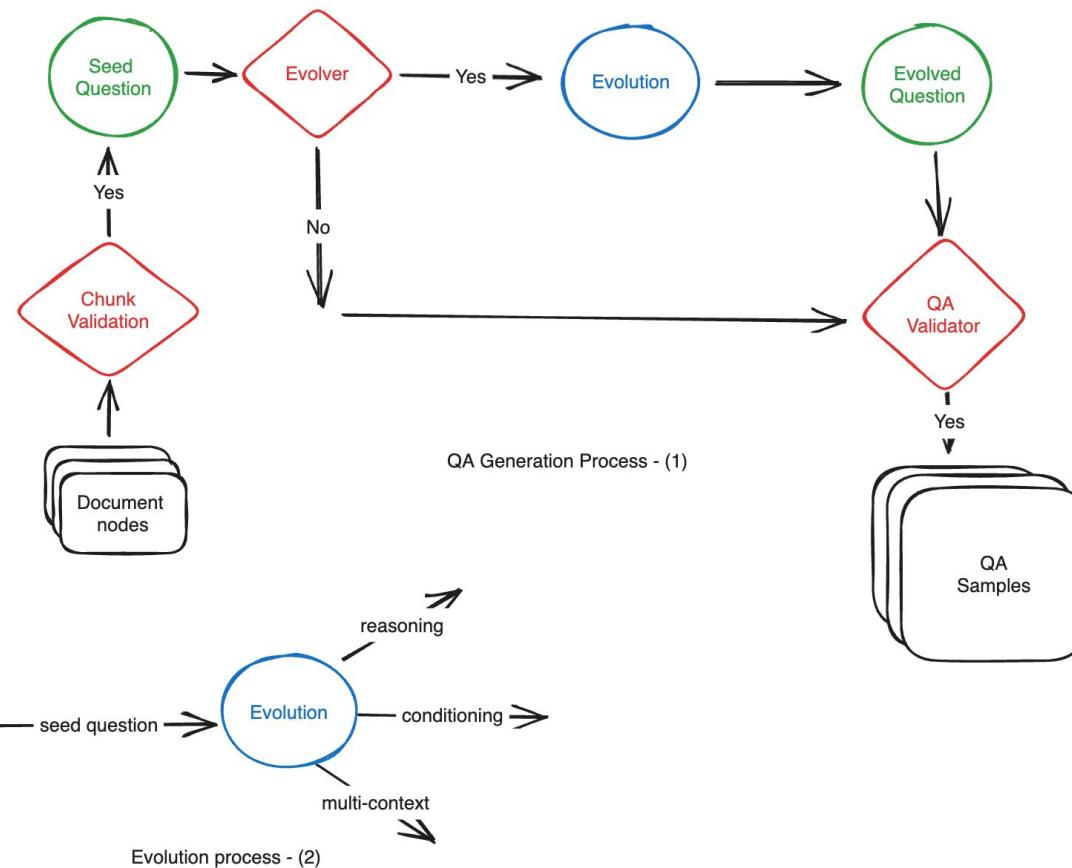
Fill-Mask • Updated Feb 19 • ↓ 26.6k • ❤ 2

▼ Expand 15 models

The TARDIS Operational Manual



RAGAS Question Generation



The TARDIS Operational Manual

Generated Questions Sample

What is the purpose of the Dimensional Stabilization Field in the TARDIS?

What steps can Time Lords take to troubleshoot time distortions in their TARDIS?

What is the significance of the Control Room in the operation of a TARDIS?

Table 2: Examples of the questions generated for different document contexts.



Next Sentence Prediction with BERT as a Dynamic Chunking Mechanism for Retrieval-Augmented Generation Systems

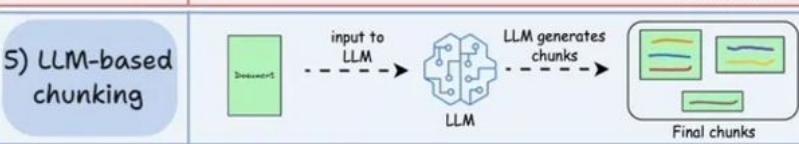
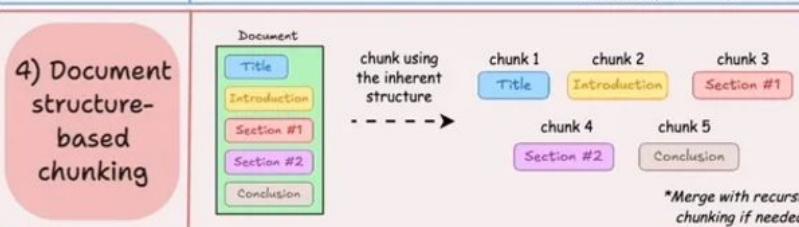
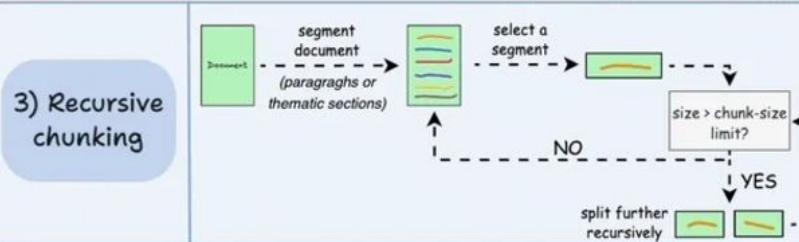
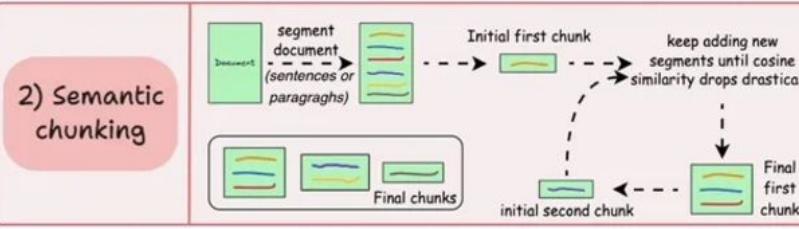
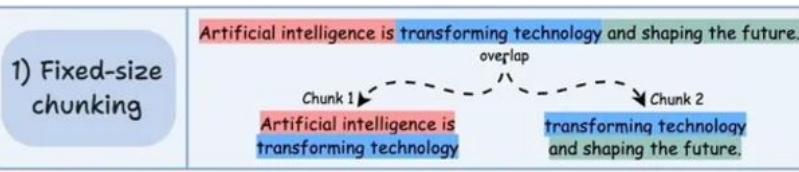
UFPel - Federal University of Pelotas

CDTec - Technology Development Center Department

H2IA - AI Development Hub

Experimental Setup - Dataset

id string	title string	context string	question string	answers sequence
56be85543aeaaa14008c9063	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɔːnseɪ/ bee-YON-say) (born September 4, 1981) is an American singer, songwriter, record producer and actress. Born and raised in Houston, Texas, she performed in various singing and dancing competitions as a child, and rose to fame in the late 1990s as lead singer of R&B girl-group Destiny's Child. Managed by her father, Mathew Knowles, the group became one of the world's best-selling girl groups of all time. Their hiatus saw the release of Beyoncé's debut album, Dangerously in Love (2003), which established her as a solo artist worldwide, earned five Grammy Awards and featured the Billboard Hot 100 number-one singles "Crazy in Love" and "Baby Boy".	When did Beyonce start becoming popular?	{ "text": ["in the late 1990s"], "answer_start": [269] }



Results - Chunk Size Distribution

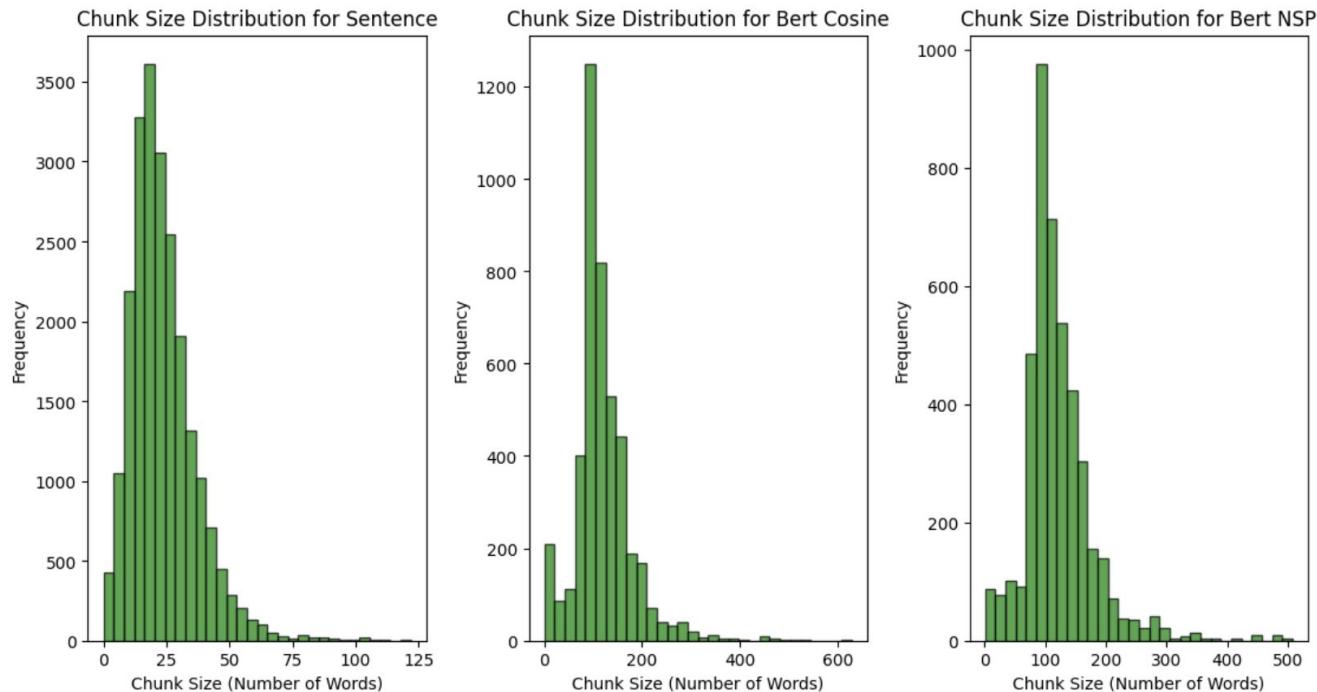


Figure 1: Chunk size frequency distributions for Sentence Chunking, Semantic Chunking using BERT Embeddings using Cosine Similarity, and Semantic Chunking using BERT for Next Sentence Prediction, respectively. Source: The authors.

Results - Experimental Results

Table 2: Test Fold Metric Results.

Approaches	Recall@3	Precision@3	Semantic-Precision@3	Processing Time
Sentence	0.9371	0.0547	0.2994	3 min
BERT Cosine	0.9829	0.0340	0.3198	44 min
<i>BERT NSP (Ours)</i>	0.9834	0.0344	0.3211	27 min



Assessing DeepSeek-R1's Performance on Brazil's Defining National Education Benchmark

UFPel - Federal University of Pelotas
CDTec - Technology Development Center Department
H2IA - AI Development Hub

Authors: Alexandre T. Bender, Gabriel A. Gomes, Ulisses B. Corrêa, Ricardo M. Araujo

Problem Statement



- LLMs show strong performance on global benchmarks but often lack evaluation in culturally-specific contexts
- Brazil's ENEM exam: multidisciplinary, culturally grounded, and decisive for university admission
- Key research question: Can reasoning-focused LLMs like DeepSeek-R1 handle ENEM's linguistic and cultural challenges?

Dataset Question Example

Discurso de Vladimir Putin sobre a Crimeia

Colegas, na mente e no coração do povo, a Crimeia sempre foi uma porção inseparável da Rússia. Essa firme convicção se baseia na verdade e na justiça e foi passada de geração em geração, ao longo do tempo, sob quaisquer circunstâncias, apesar de todas as drásticas mudanças que nosso país atravessou durante todo o século 20.

Considerando a dinâmica geopolítica subjacente ao texto, a justificativa utilizada por Vladimir Putin, em 2014, para anexação dessa península apela para o argumento de que

- A) “as populações com idioma comum devem estar submetidas à mesma autoridade estatal.”
- B) “o imperialismo soviético havia se acomodado às pretensões das potências vizinhas.”
- C) “os organismos transnacionais são incapazes de solucionar disputas territoriais.”
- D) “a integração regional supõe a livre circulação de pessoas e mercadorias.”
- E) “a expulsão das forças navais ocidentais garantiria a soberania nacional.”

Chain-of-Thought Example

<think>

Okay, so I need to figure out the answer to this question about Vladimir Putin's justification for annexing Crimea in 2014. The question is in Portuguese, but I'll try to understand it step by step.

The text says that in the minds and hearts of the people, Crimea has always been an inseparable part of Russia. This conviction is based on truth and justice and was passed down through generations despite all the drastic changes the country went through during the 20th century. So the main idea here seems to be about historical and cultural ties.

Now, looking at the question: it's asking which argument Putin used in 2014 for annexing Crimea. The options are A to E.

Let me think about what I know regarding this situation. In 2014, Russia annexed Crimea citing several reasons. One of them was the protection of ethnic Russians and their cultural ties. They argued that Crimea had historical connections with Russia and that the people there wanted to be part of Russia again. There were also issues with the government in Kiev at the time, which was seen as unstable or not representing Russian-speaking populations.

Chain-of-Thought Example

Looking at the options:

A) Populações com idioma comum devem estar submetidas à mesma autoridade estatal.

Translation: “Populations with a common language should be subject to the same state authority.” So this is about linguistic unity leading to same governance. That’s similar to the idea of a nation-state based on language.

B) O imperialismo soviético havia se acomodado às pretensões das potências vizinhas.

Translation: “Soviet imperialism had accommodated the claims of neighboring powers.” I’m not sure about this one. It seems more about Soviet policies towards neighbors, which doesn’t directly tie to Crimea’s annexation justification.

C) Os organismos transnacionais são incapazes de solucionar disputas territoriais.

Translation: “Transnational organizations are unable to solve territorial disputes.” This is talking about international bodies failing. I don’t recall Putin specifically citing this as a reason, but more about the internal aspects of Crimea.

Chain-of-Thought Example

D) A integração regional supõe a livre circulação de pessoas e mercadorias.

Translation: ‘Regional integration implies free circulation of people and goods.’ That sounds like economic or movement-based reasons, which doesn’t directly address why annex Crimea.

E) A expulsão das forças navais ocidentais garantiria a soberania nacional.

Translation: ‘The expulsion of Western naval forces would ensure national sovereignty.’ This seems related to military presence. I think Russia was concerned about NATO expanding and having bases near them, but I’m not sure if that was the main argument in 2014.

Chain-of-Thought Example

Going back, the text mentions the conviction based on truth and justice passed through generations. So it's more about historical and cultural ties rather than language or military presence. But looking at option A, it talks about common language leading to same authority. That could tie into the idea that since they share a language (Russian), they should be under Russian rule.

Wait, but in 2014, the main argument was more about protecting ethnic Russians and historical ties rather than just language. Also, another angle is the idea of self-determination. The people in Crimea wanted to join Russia because of their cultural and historical connections.

But looking at the options again, option A seems to focus on language leading to same state authority. So if a population speaks Russian, they should be part of Russia. That could fit with Putin's argument that Crimeans are Russians and thus should be under Russian governance.

Chain-of-Thought Example

Alternatively, maybe the text in the question is more about cultural/ethnic identity rather than just language. But the options don't have an option that directly addresses ethnicity; only A talks about language leading to same authority.

So I'm leaning towards option A as the answer because it's about linguistic unity leading to state authority, which ties into the argument of shared history and identity through language.

</think>

The correct answer is A.

What is Pass@k?

$$\text{Pass}@k(n_i, c_i, k) = 1 - \frac{\binom{n_i - c_i}{k}}{\binom{n_i}{k}}$$

- Evaluates LLMs with multiple attempts allowed.
- Checks if a correct solution appears within k samples.
- Pass@1: model gets correct answer on first try.
- Pass@k: correct answer found within k attempts.



$$\text{Pass@k}(n_i, c_i, k) = 1 - \frac{\binom{n_i - c_i}{k}}{\binom{n_i}{k}}$$

Example

Suppose that for a given problem you generated $n_i = 5$ responses, and $c_i = 2$ of them are correct.

We want to calculate **Pass@3** (the probability of having at least one correct answer among 3 samples):

$$\text{Pass@3} = 1 - \frac{\binom{3}{3}}{\binom{5}{3}} = 1 - \frac{1}{10} = 0.9$$

There is a **90% chance** of obtaining at least one correct response if we randomly pick 3 generations.

Table 2. Pass@k Evaluation Results by Year and Subject Area

Year	Area	pass@1	pass@2	pass@3	pass@4	pass@5
2022	Human Sciences	0.922	0.933	0.938	0.941	0.944
	Languages and Codes	0.867	0.927	0.952	0.963	0.969
	Mathematics	0.730	0.799	0.822	0.836	0.846
	Natural Sciences	0.763	0.827	0.852	0.866	0.875
2023	Human Sciences	0.990	0.999	1.000	1.000	1.000
	Languages and Codes	0.895	0.921	0.932	0.938	0.942
	Mathematics	0.620	0.700	0.737	0.759	0.776
	Natural Sciences	0.881	0.930	0.948	0.960	0.970
2024	Human Sciences	0.988	0.999	1.000	1.000	1.000
	Languages and Codes	0.862	0.900	0.911	0.916	0.918
	Mathematics	0.885	0.920	0.925	0.926	0.926
	Natural Sciences	0.844	0.888	0.914	0.930	0.938

Table 2. Pass@k Evaluation Results by Year and Subject Area

Year	Area	pass@1	pass@2	pass@3	pass@4	pass@5
2022	Human Sciences	0.922	0.933	0.938	0.941	0.944
	Languages and Codes	0.867	0.927	0.952	0.963	0.969
	Mathematics	0.730	0.799	0.822	0.836	0.846
	Natural Sciences	0.763	0.827	0.852	0.866	0.875
2023	Human Sciences	0.990	0.999	1.000	1.000	1.000
	Languages and Codes	0.895	0.921	0.932	0.938	0.942
	Mathematics	0.620	0.700	0.737	0.759	0.776
	Natural Sciences	0.881	0.930	0.948	0.960	0.970
2024	Human Sciences	0.988	0.999	1.000	1.000	1.000
	Languages and Codes	0.862	0.900	0.911	0.916	0.918
	Mathematics	0.885	0.920	0.925	0.926	0.926
	Natural Sciences	0.844	0.888	0.914	0.930	0.938

Table 2. Pass@k Evaluation Results by Year and Subject Area

	Year Area	pass@1	pass@2	pass@3	pass@4	pass@5
2022	Human Sciences	0.922	0.933	0.938	0.941	0.944
	Languages and Codes	0.867	0.927	0.952	0.963	0.969
	Mathematics	0.730	0.799	0.822	0.836	0.846
	Natural Sciences	0.763	0.827	0.852	0.866	0.875
2023	Human Sciences	0.990	0.999	1.000	1.000	1.000
	Languages and Codes	0.895	0.921	0.932	0.938	0.942
	Mathematics	0.620	0.700	0.737	0.759	0.776
	Natural Sciences	0.881	0.930	0.948	0.960	0.970
2024	Human Sciences	0.988	0.999	1.000	1.000	1.000
	Languages and Codes	0.862	0.900	0.911	0.916	0.918
	Mathematics	0.885	0.920	0.925	0.926	0.926
	Natural Sciences	0.844	0.888	0.914	0.930	0.938

Post-Training

Where do LLMs come from?

Well, we train them

LLM Training

LLM Training

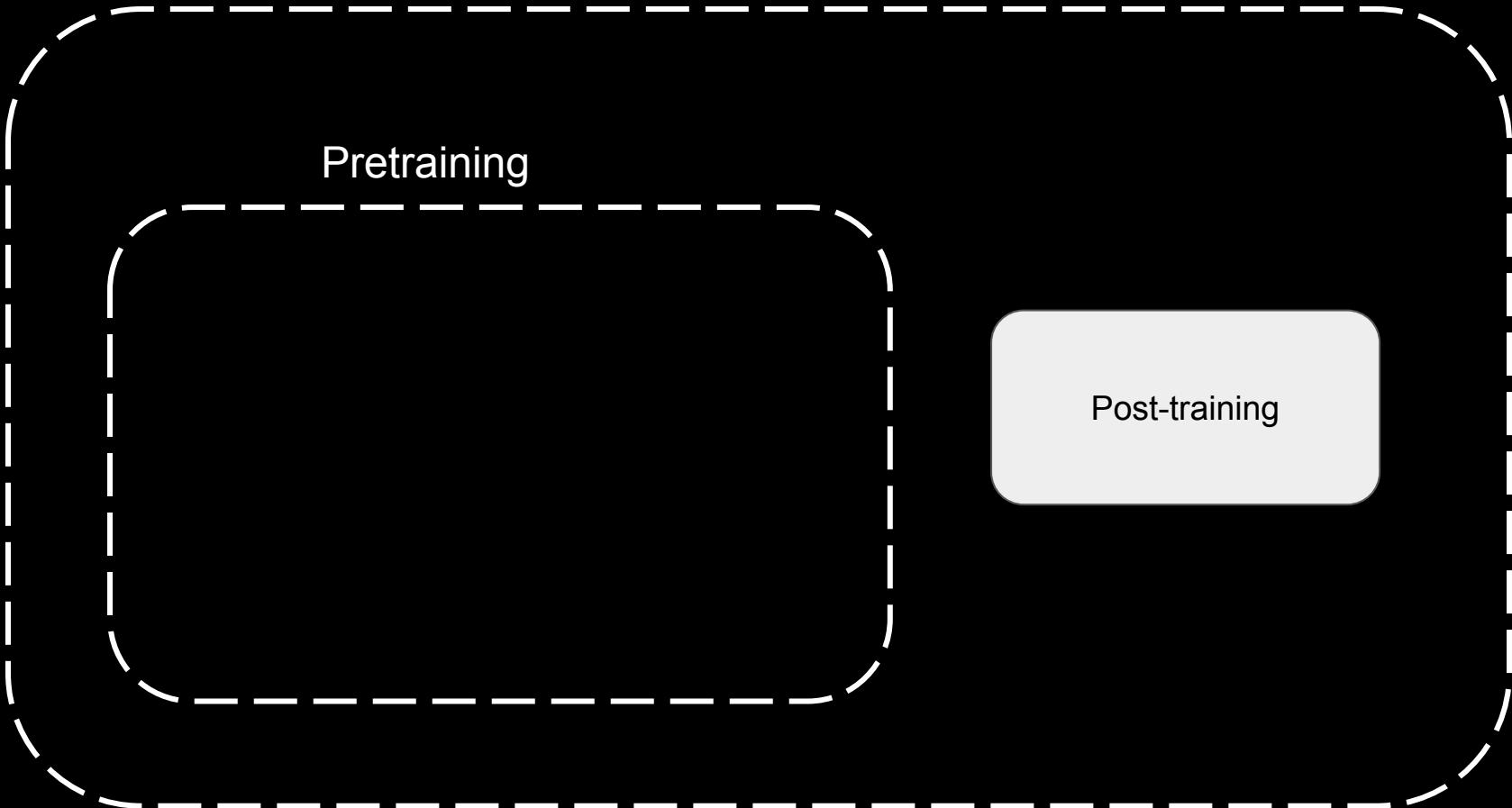


LLM Training

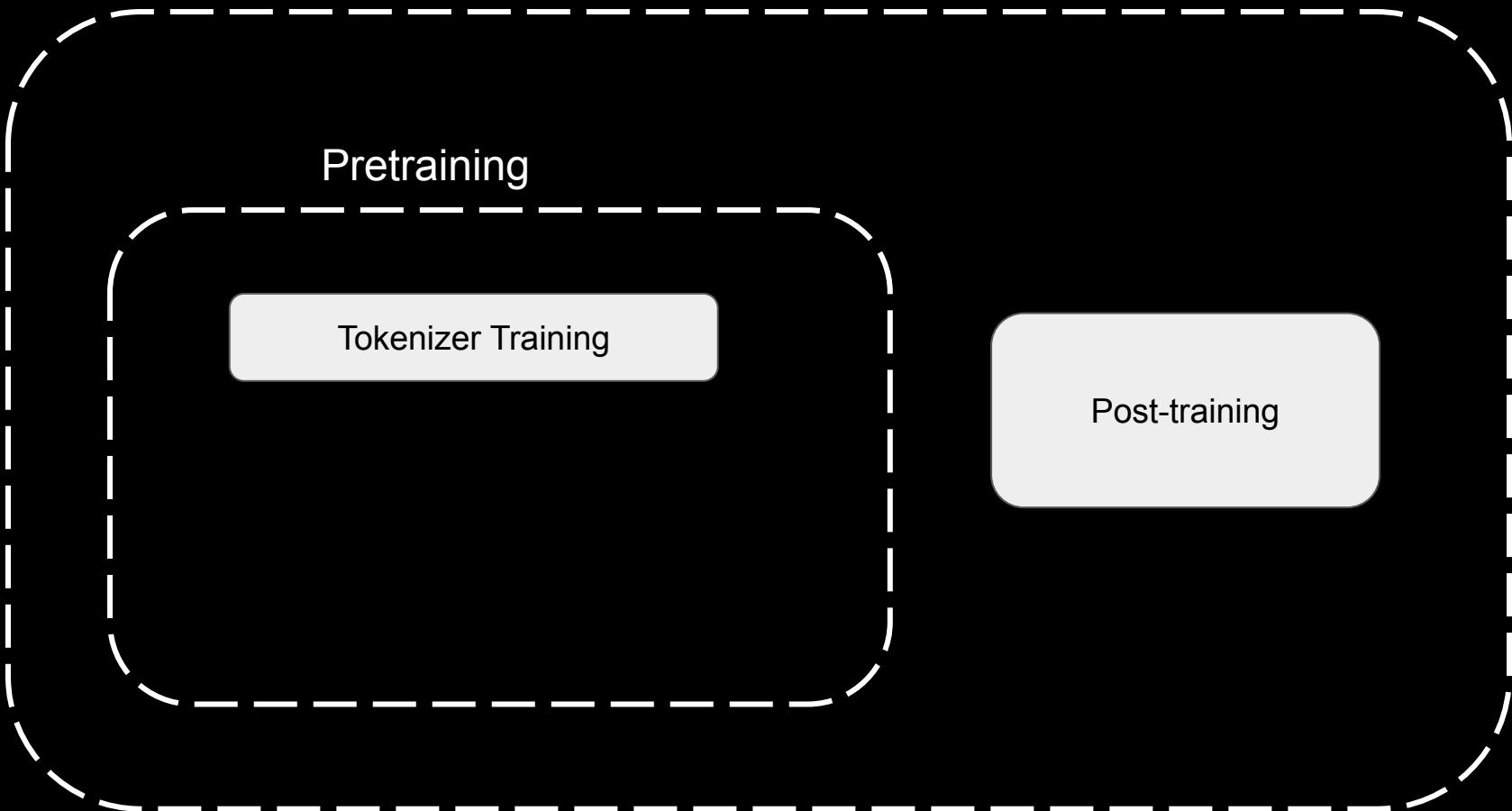
Pretraining

Post-training

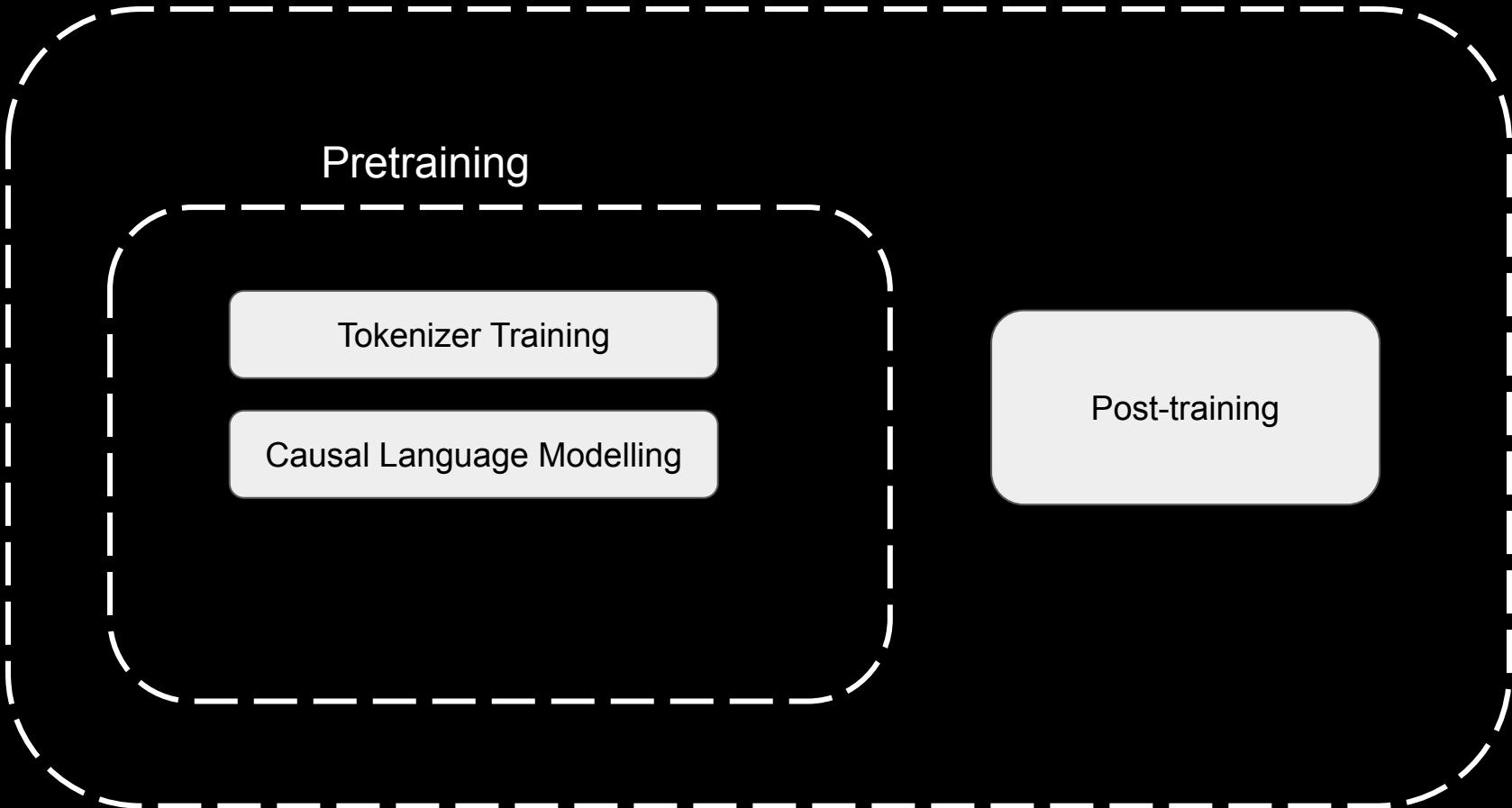
LLM Training



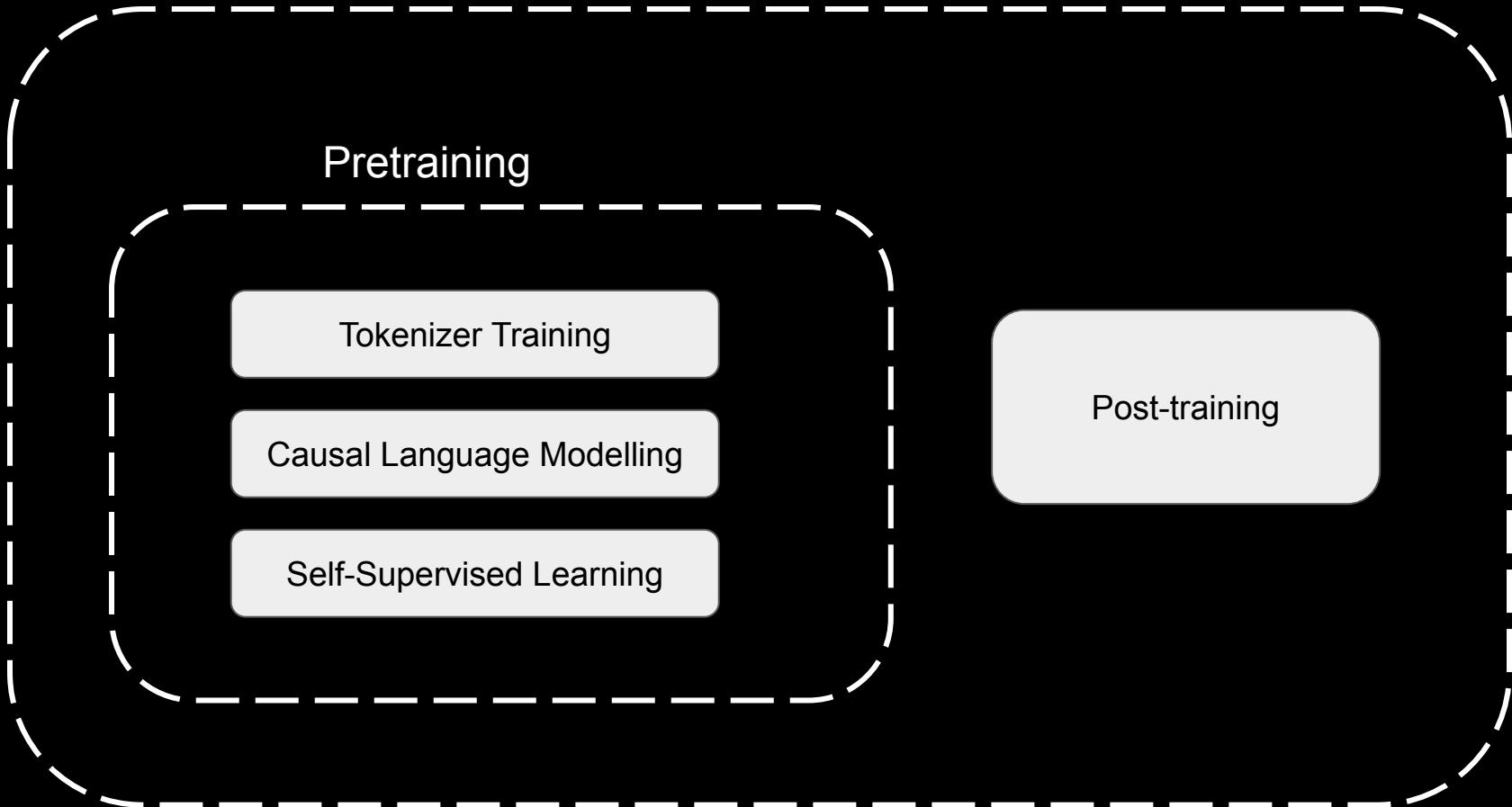
LLM Training



LLM Training



LLM Training



LLM Training

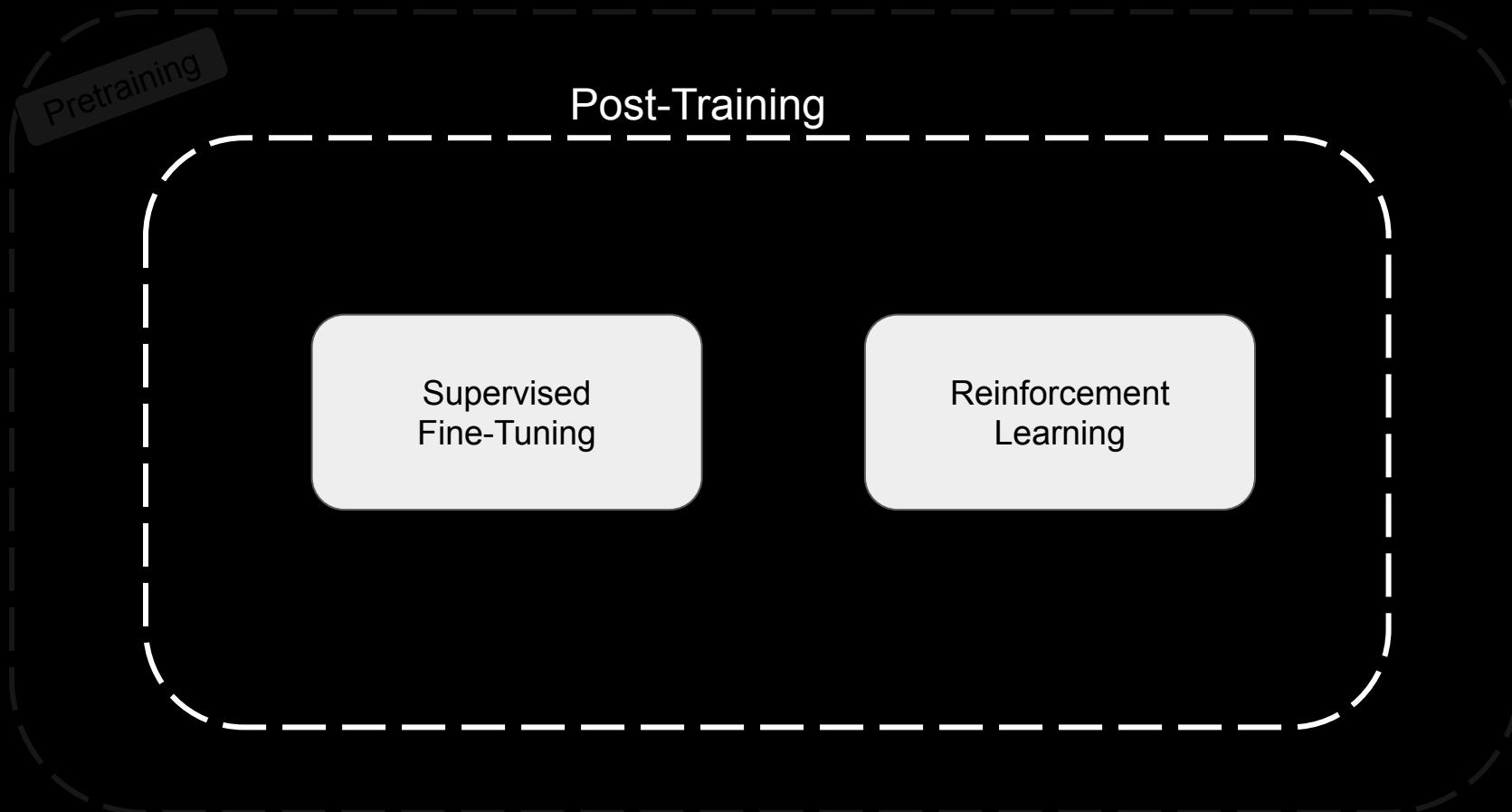
Pretraining

Post-training

LLM Training



LLM Training



LLM Training

Pretraining

Post-Training

Supervised Fine-Tuning

Reinforcement
Learning

LLM Training

Pretraining

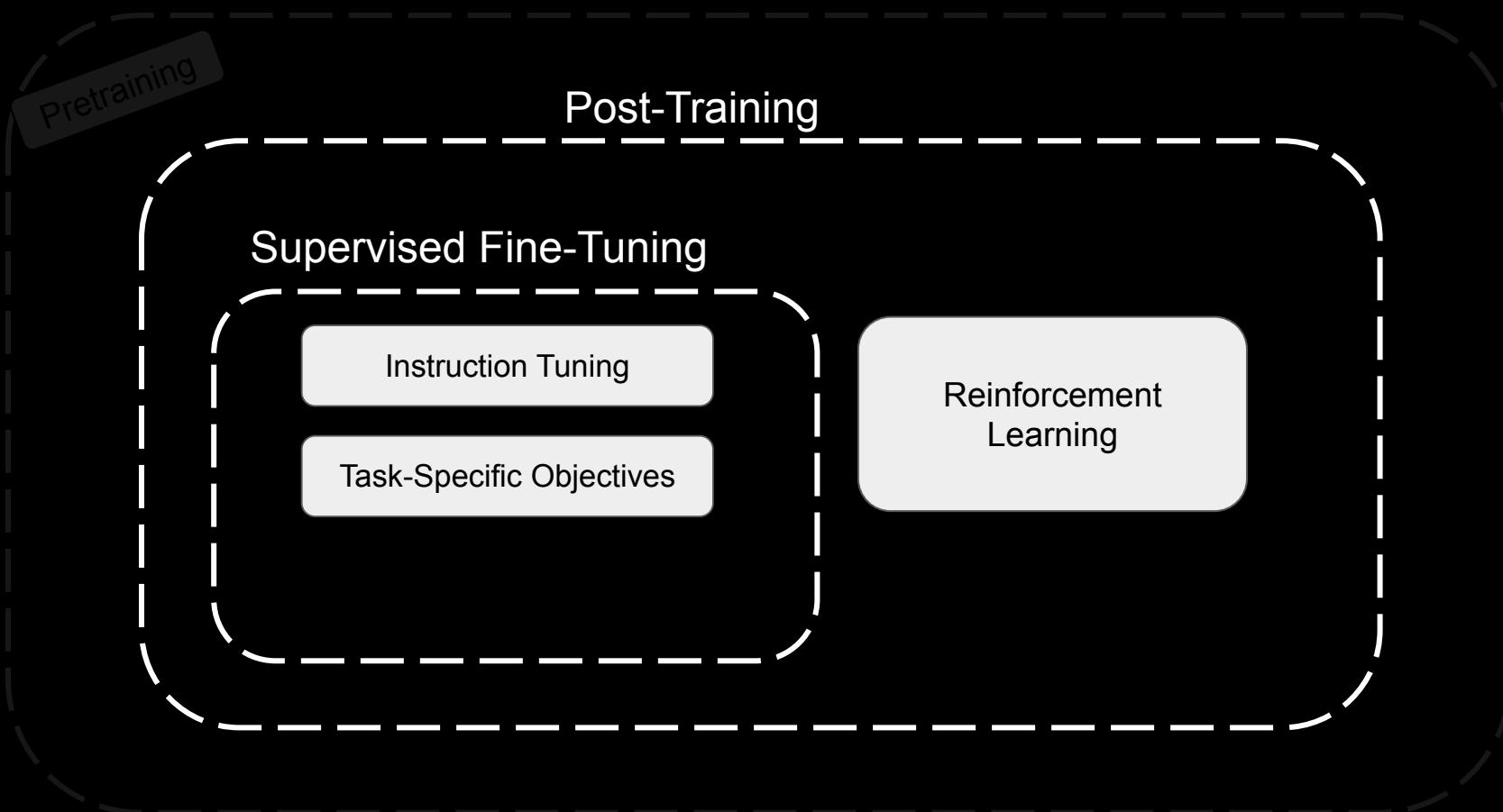
Post-Training

Supervised Fine-Tuning

Instruction Tuning

Reinforcement
Learning

LLM Training



LLM Training

Pretraining

Post-Training

Supervised Fine-Tuning

Instruction Tuning

Task-Specific Objectives

Supervised Learning

Reinforcement
Learning

LLM Training

Pretraining

Post-Training

SFT

Reinforcement Learning

LLM Training

Pretraining

Post-Training

Reinforcement Learning

SFT

Align Model with Human Preferences

LLM Training

Pretraining

Post-Training

SFT

Reinforcement Learning

Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

LLM Training

Pretraining

Post-Training

SFT

Reinforcement Learning

Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

LLM Training

Pretraining

Post-Training

SFT

Reinforcement Learning

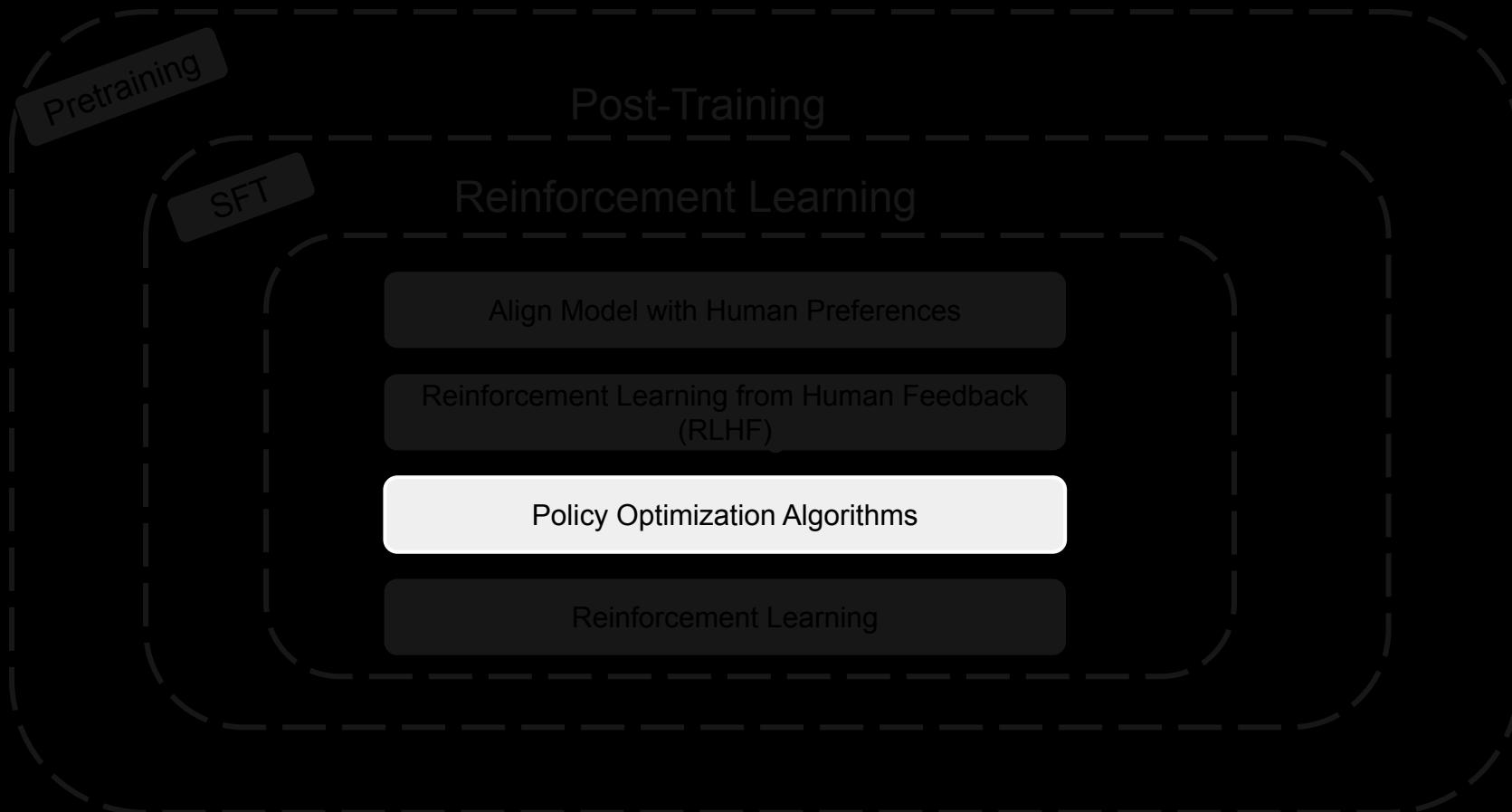
Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

Reinforcement Learning

LLM Training



LLM Training

Proximal Policy
Optimization
(PPO)

Post-Training

Reinforcement Learning

Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

Reinforcement Learning

Proximal Policy
Optimization
(PPO)

Direct Preference
Optimization (DPO)

Reinforcement Learning

Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

Reinforcement Learning

Proximal Policy
Optimization
(PPO)

Direct Preference
Optimization (DPO)

Group Relative Policy
Optimization (GRPO)

Reinforcement Learning

Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

Reinforcement Learning

Proximal Policy
Optimization
(PPO)
2017

Direct Preference
Optimization (DPO)

2023

Group Relative Policy
Optimization (GRPO)

2024

Reinforcement Learning

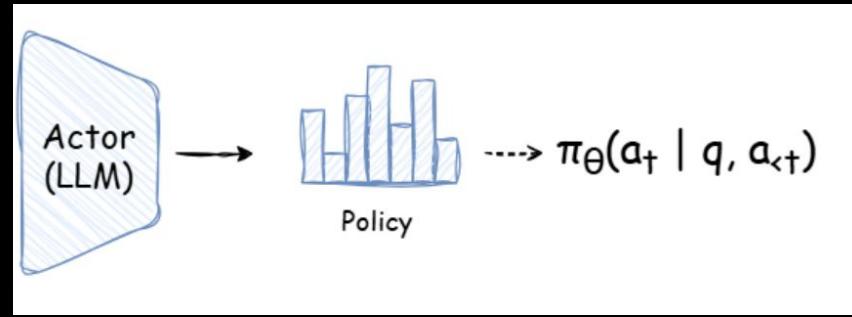
Align Model with Human Preferences

Reinforcement Learning from Human Feedback
(RLHF)

Policy Optimization Algorithms

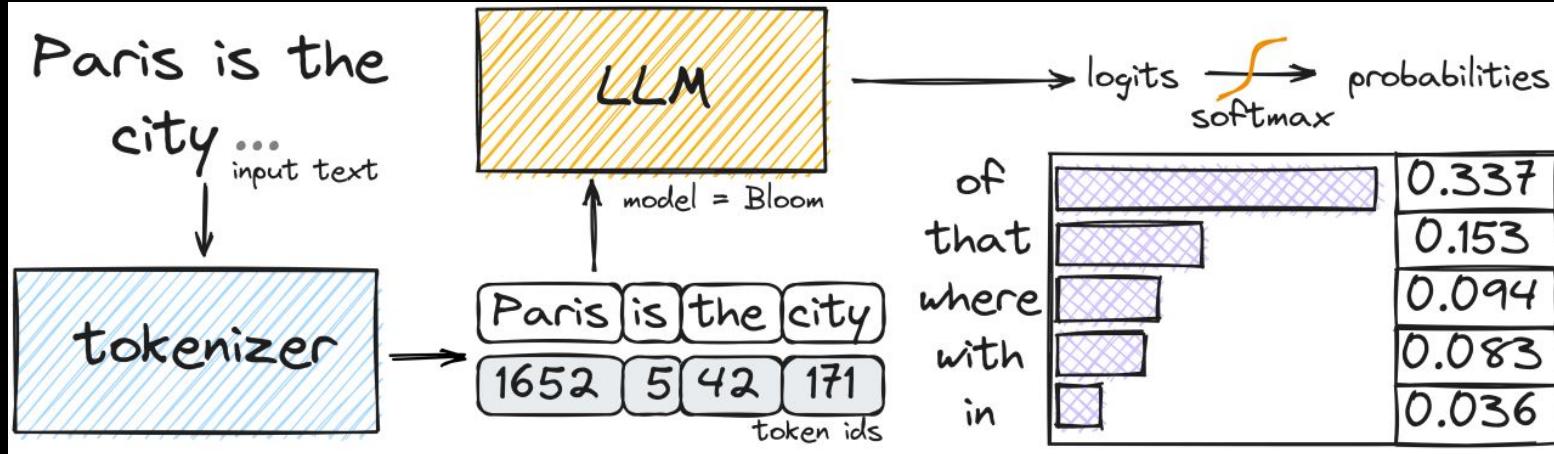
Reinforcement Learning

Wtf is a policy?



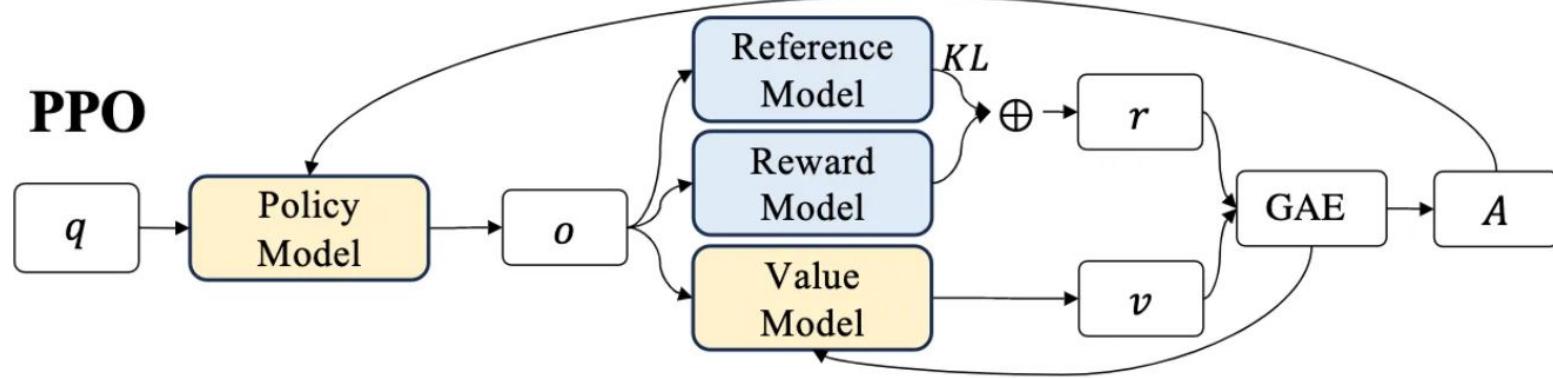
Wtf is a policy?

Fancy word for model being aligned



LLMs as output distribution functions

$$f(\langle \text{start} \rangle) = \begin{pmatrix} P(\text{The}) = .5 \\ P(\text{cat}) = .1 \\ P(\text{is}) = .2 \\ P(\text{blue}) = .05 \\ P(\text{black}) = .1 \\ P(\text{green}) = .04 \\ P(\langle \text{end} \rangle) = .01 \end{pmatrix}$$

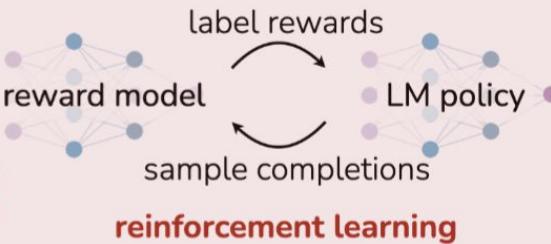


Proximal Policy Optimization (PPO)

Reward models are hard

Reinforcement Learning from Human Feedback (RLHF)

x: "write me a poem about
the history of jazz"



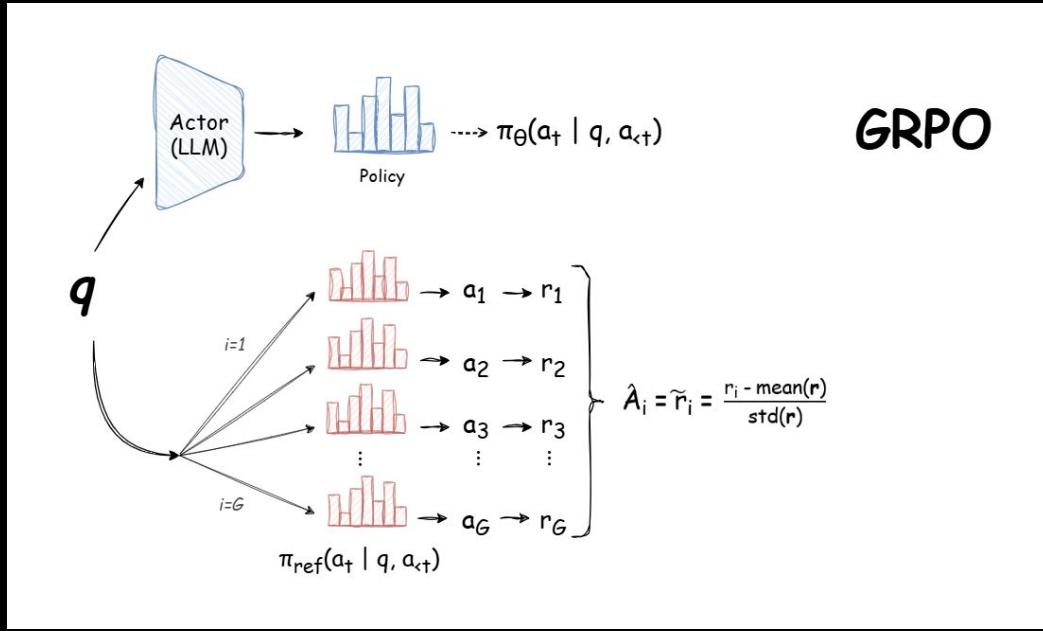
Direct Preference Optimization (DPO)

x: "write me a poem about
the history of jazz"



Direct Preference Optimization (DPO)

Your model is already a reward model

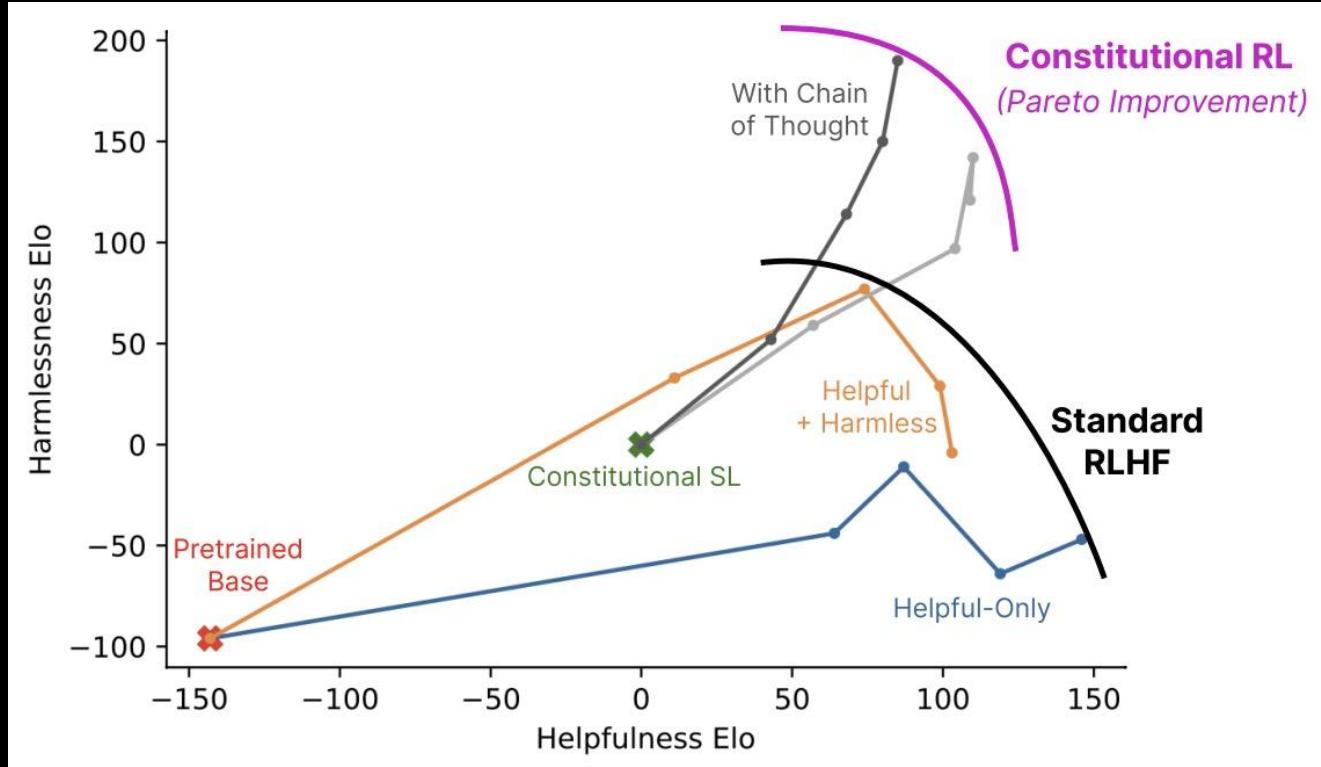


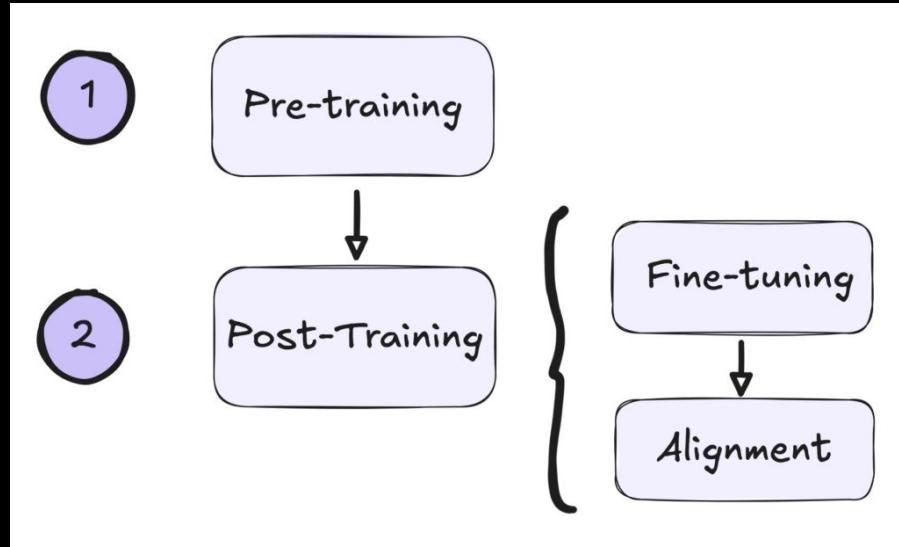
Group Relative Policy Optimization (GRPO)

Reward function soup

Challenges

I'M SORRY,
DAVE.
I'M AFRAID I
CAN'T
DO THAT.





Where to add new behaviors to a model?

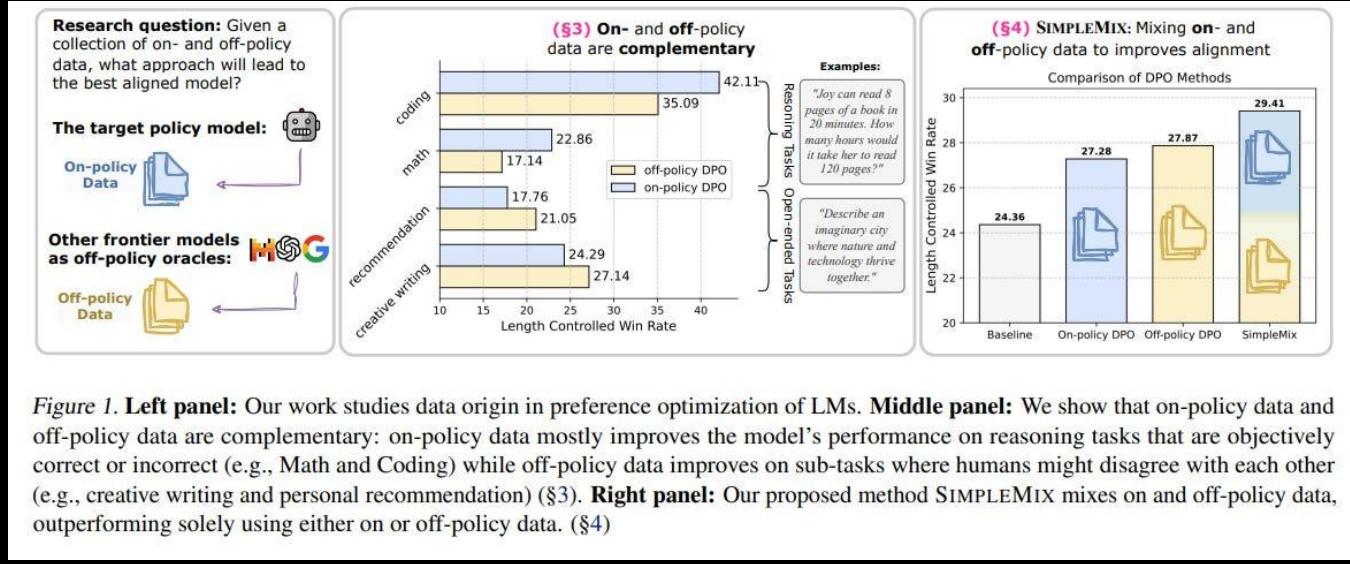


Figure 1. Left panel: Our work studies data origin in preference optimization of LMs. **Middle panel:** We show that on-policy data and off-policy data are complementary: on-policy data mostly improves the model's performance on reasoning tasks that are objectively correct or incorrect (e.g., Math and Coding) while off-policy data improves on sub-tasks where humans might disagree with each other (e.g., creative writing and personal recommendation) (§3). **Right panel:** Our proposed method SIMPLEMIX mixes on and off-policy data, outperforming solely using either on or off-policy data. (§4)

On-Policy vs Off-Policy

Does the data we're training with comes from our current policy or some other?

