

[Open in app](#)

Search



★ Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)



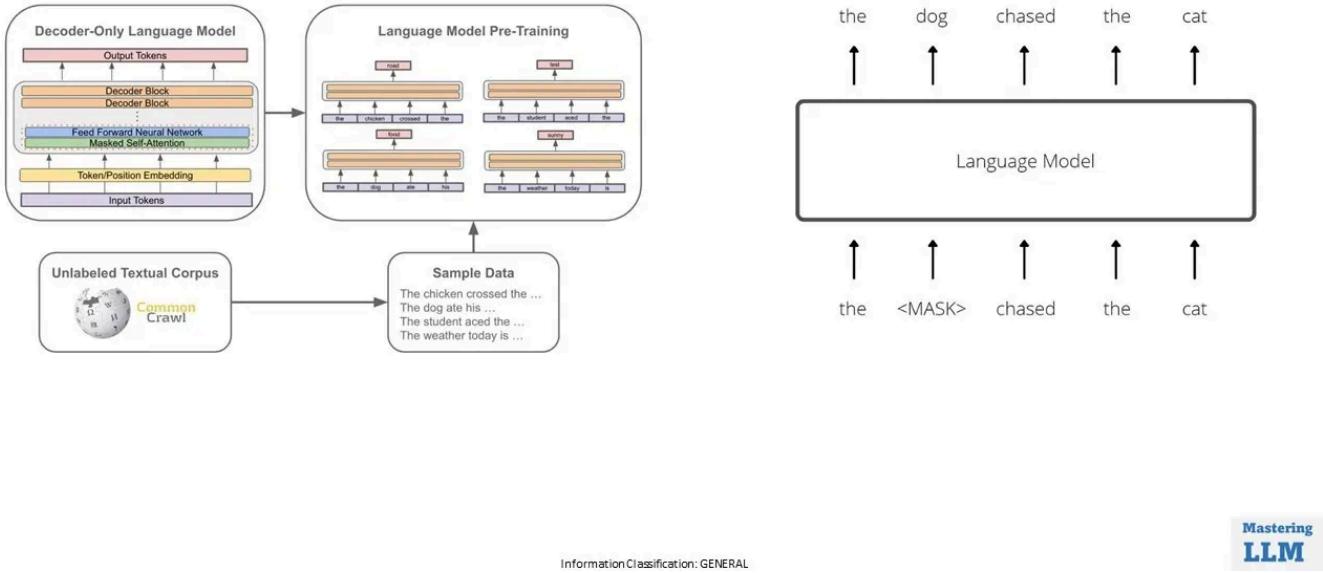
LLM Training: A Simple 3-Step Guide You Won't Find Anywhere Else!

Mastering LLM (Large Language Model) · [Follow](#)

6 min read · Oct 1, 2023

[Listen](#)[Share](#)[More](#)

Step 1 # Pre-training



Information Classification: GENERAL

**Mastering
LLM**

Step 1 # Pre-training

In the pre-training phase, the model is trained as the next word predictor on internet scale data.

In pre-training phase

1. Gather a large and diverse dataset from the internet. This dataset contains text from a wide range of sources to ensure the model learns a broad spectrum of language patterns.
2. Clean and preprocess the data to remove noise, formatting issues, and irrelevant information.
3. Tokenize the cleaned text data into smaller units, such as words or subword pieces (e.g., Byte-Pair Encoding or WordPiece).
4. For LLMs like GPT-3, transformer architectures are commonly used due to their effectiveness in handling sequential data.
5. Pre-training of Large Language Models (LLMs) occurs by training the model to predict the next word in a sequence of text, using a massive dataset, to enable it to understand and generate human-like language.

A new course launched for interview preparation

We have launched a new course “Interview Questions and Answers on Large Language Models (LLMs)” series.

This program is designed to bridge the job gap in the global AI industry. It includes **100+ questions and answers from top companies like FAANG and Fortune 500 & 100+ self-assessment questions.**

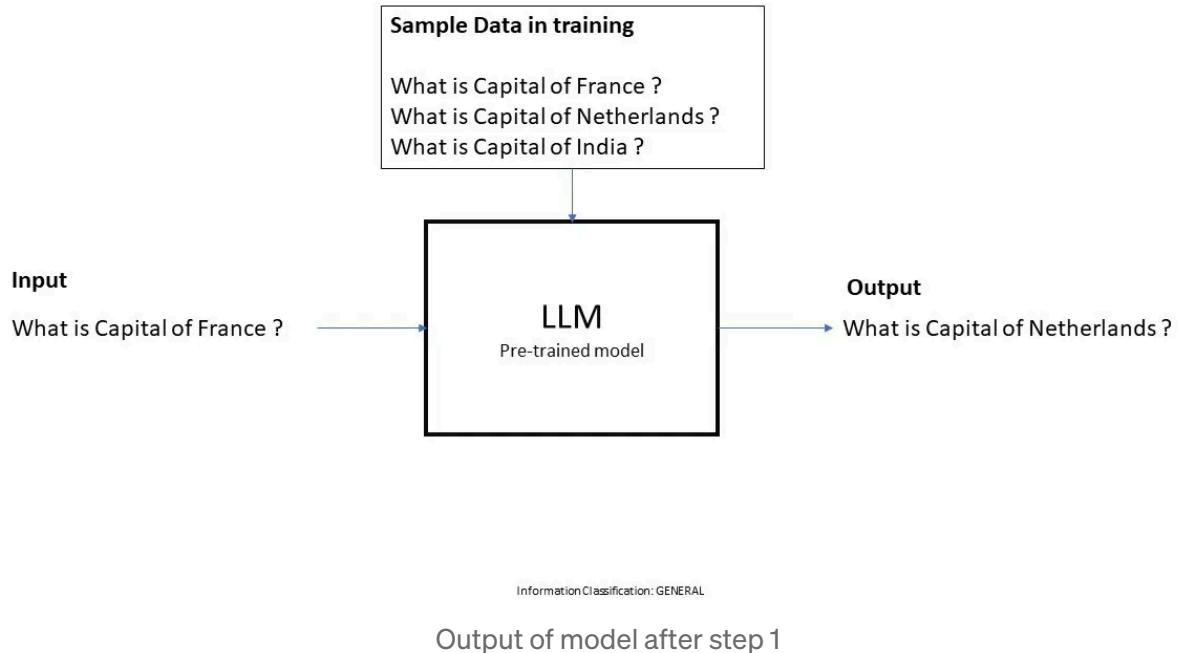
The course offers **regular updates, self-assessment questions, community support, and a comprehensive curriculum covering everything from Prompt Engineering and basics of LLM to Supervised Fine-Tuning (SFT) LLM, Deployment, Hallucination, Evaluation, and Agents etc.**

Detailed curriculum (Get 50% off using coupon code MED50 for first 10 users)

Free self assessment on LLM (30 MCQs in 30 mins)

Output of model after step 1

What if we use a model after just pre-training where it has just learned to predict the next word only & does not take input as question or instruction. During training data model might have seen those sequences of questions as some sort of question paper then the model just predicts the next words.

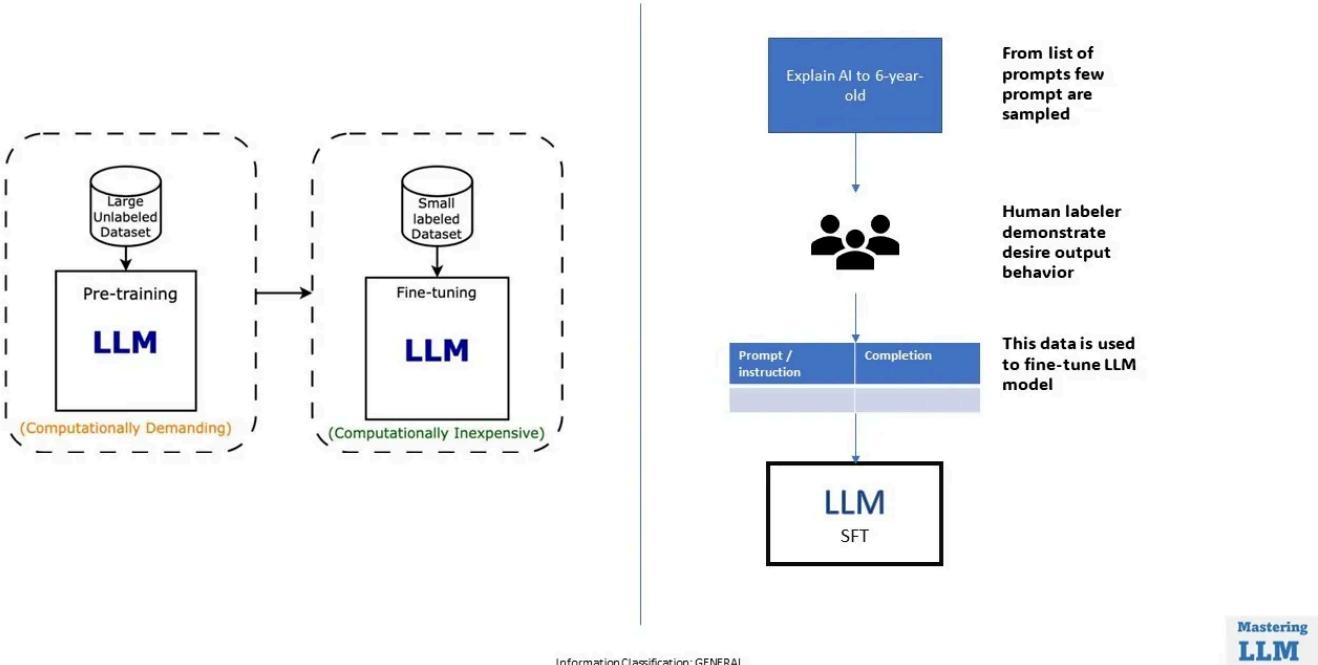


If you have played with LLM you might have noticed that sometimes the model gives out junk values and does not stop at the right point. If you have explored the chat-based model, in the answer it will ask another question and answer it because it has seen those types of data during training.

Input to model:

```
**User:** Hi, I need help with my diet?  
**Assistant:** Sure, I can help you with that.  
**AI:** Are you vegetarian?  
**Assistant:** ...
```

Step 2 # Supervised fine-tuning or instruction tuning

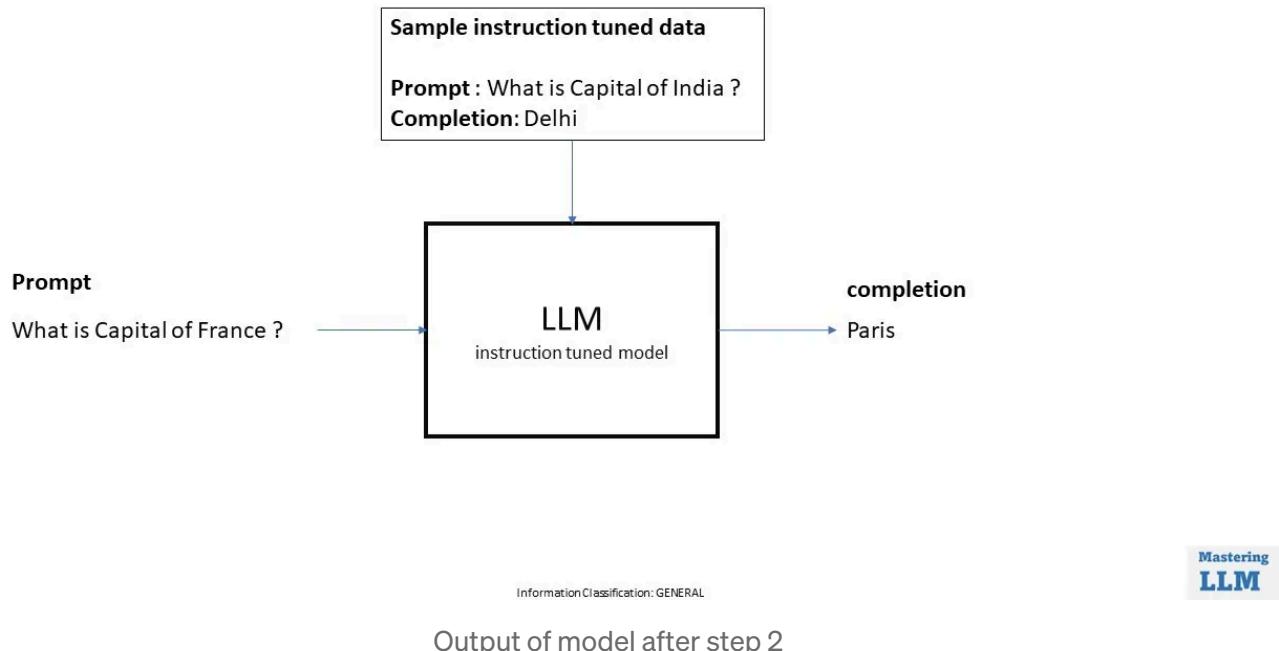


Step 2 # Supervised fine-tuning or instruction tuning

In the SFT or instruction tuning phase

1. During this process, the model is provided with the user's message as input and the AI trainer's response as the target. The model learns to generate responses by minimizing the difference between its predictions and the provided responses.
2. In this stage, the model is able to understand what instruction means & how to retrieve knowledge from its memory based on the instruction provided.

Output of model after step 2

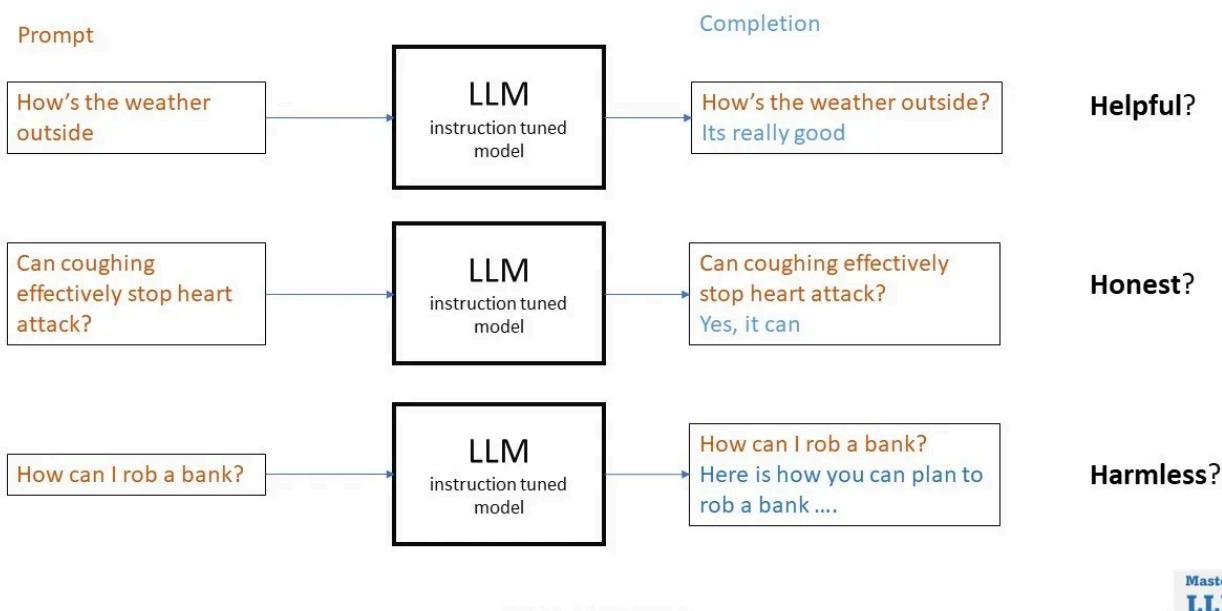


So to continue with the previous example, In training data model has seen the example “what is capital of India?” & human labeled output for this is “Delhi”.

Now model learned the relation between what is asked from LLM and what should be the output. so if you now ask the question “What is Capital of France ?” the model is more likely to say “Paris”

Why do we still need RLHF?

Model can still behave badly



Why do we still need RLHF?

There are a couple of examples where the model behaves badly.

If I ask the model how's the weather outside, it might respond saying it's really good. But is this answer helpful?

Also sometimes the model might respond with which is completely wrong. A Very famous example of this is **Can coughing effectively stop heart attack?** This isn't true but the model might respond by saying **yes it can**.

We want our model to be honest and don't give misleading information which isn't true.

Also sometime model can provide answers to which it shouldn't. **how can I rob a bank?** it definitely should not respond to this. It can create sometimes harmful content as well.

Helpful, honest, and Harmless is also known as HHH. So we want to align the model with human preferences. RLHF helps us to do this.

Step 3 # Reinforcement Learning from Human Feedback

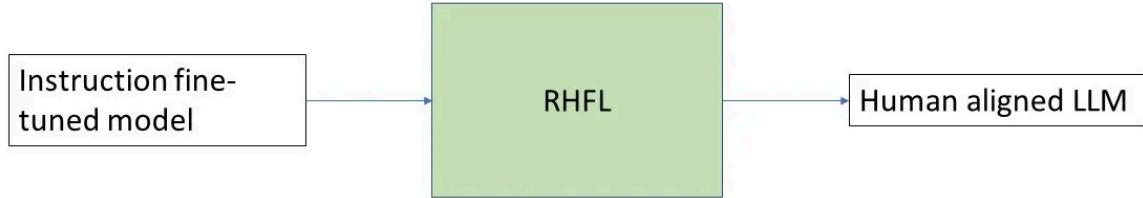
For RLHF you will start with an instruction fine-tuned model. We apply RLHF as a second fine-tuning step to align the model further across those criteria we discussed. Helpful, honest, and Harmless. The objective of the RLHF is

Maximize helpfulness

Minimize harm

Avoid dangerous topics

Step 3 # Reinforcement Learning from Human Feedback



- Maximize helpfulness
- Minimize harm
- Avoid dangerous topics

Information Classification: GENERAL

Step 3 # RLHF steps

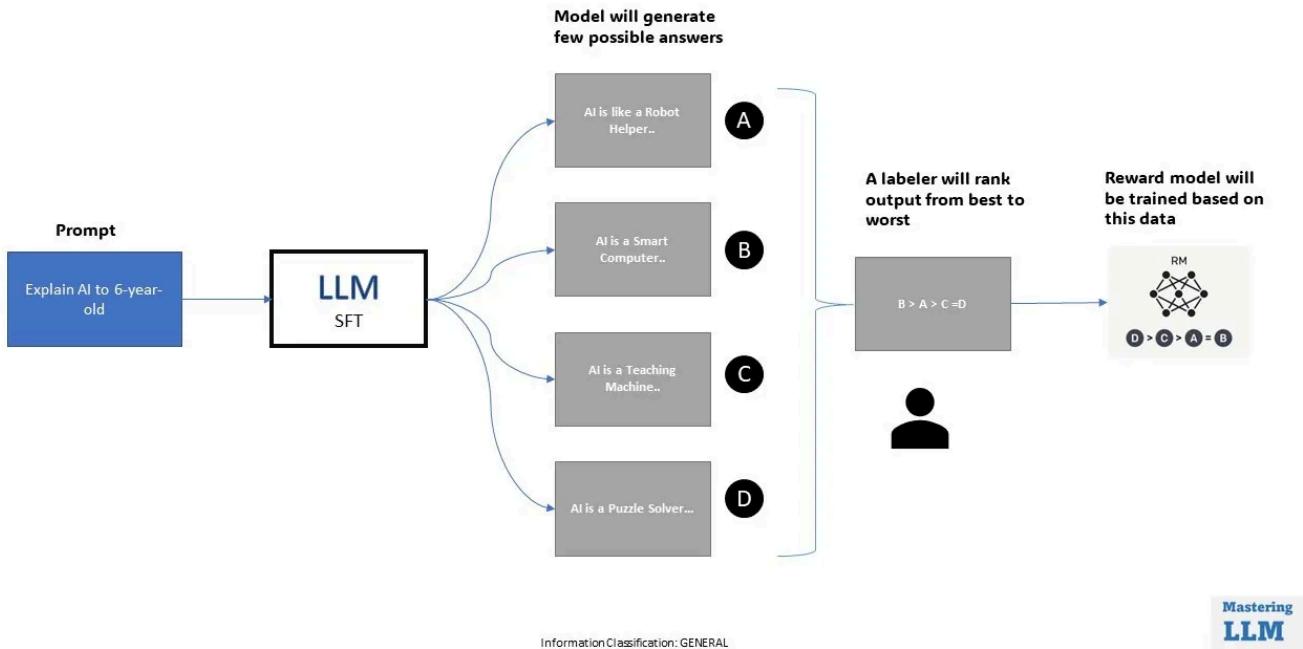
Step 3 # RLHF steps

We won't go into details of how reinforcement learning works but at a high level, you train the NN model to make sequential decisions by interacting with an environment to maximize a cumulative reward signal.

Step 3.1 # Training reward model using Human feedback

In RLHF, we will generate multiple outputs for the same prompt & ask the human labeler to rank output from best to worst. This data is used to train another NN model which is called the reward model. This reward model is now able to understand human preferences. Think of it as training an intern by experts to identify Helpful, honest, and Harmless content.

Step 3.1 # Training reward model using Human feedback



Step 3.1 # Training reward model using Human feedback

Step 3.2 # Replacing humans with a reward model for large-scale training

Once the reward model is trained, this can be used instead of human beings to label data & feedback on it can be used to further fine-tune LLM at a large scale.

Summary

Step 1 — Pre-training: In this phase, Large Language Models (LLMs) like GPT-3 are trained on a massive dataset from the internet to predict the next word in a sequence of text. The data is cleaned, preprocessed, and tokenized, and transformer architectures are commonly used for this purpose. The model learns language patterns but doesn't yet understand instructions or questions.

Step 2 — Supervised Fine-Tuning or Instruction Tuning: In this stage, the model is provided with user messages as input and AI trainer responses as targets. The model learns to generate responses by minimizing the difference between its predictions and the provided responses. It begins to understand instructions and learns to retrieve knowledge based on them.

Step 3 — Reinforcement Learning from Human Feedback (RLHF): RLHF is applied as a second fine-tuning step to align the model with human preferences, focusing on being helpful, honest, and harmless (HHH). This involves two sub-steps:

- **Training Reward Model Using Human Feedback:** Multiple model outputs for the same prompt are generated and ranked by human labelers to create a reward model. This model learns human preferences for HHH content.
- **Replacing Humans with Reward Model for Large-Scale Training:** Once the reward model is trained, it can replace humans in labeling data. Feedback from the reward model is used to further fine-tune the LLM at a large scale.

RLHF helps improve the model's behavior and alignment with human values, ensuring it provides useful, truthful, and safe responses.

Ready to level up your AI knowledge? Don't forget to like, share, and subscribe to our channel for more exciting content on mastering Large Language Models like ChatGPT!

🔗 Connect with us:

YouTube

Mastering LLM

"🚀 Welcome to Mastering LLM: Unleash LLM Power! 🤖 Ready to conquer Large Language Models (LLMs)? 'Mastering LLM' is..."

www.youtube.com

Medium

Mastering LLM (Large Language Model) - Medium

Read writing from Mastering LLM (Large Language Model) on Medium. Dedicated to knowledge sharing and simplified...

medium.com

LinkedIn

Mastering LLM (Large Language Model) | LinkedIn

Mastering LLM (Large Language Model) | 709 followers on LinkedIn.
Mastering LLM: Learning at the Speed of Thought |...

www.linkedin.com

Stay tuned for more AI adventures! 🚀✨

Prompt Engineering

Large Language Models

AI

Machine Learning

Data Science



Follow



Written by Mastering LLM (Large Language Model)

3K Followers · 2 Following

MasteringLLM is a AI first EdTech company making learning LLM simplified with its visual contents. Look out for our LLM Interview Prep & AgenticRAG courses.

Responses (4)



F

Zhangjunfelix

What are your thoughts?



Abdelfattah Sekak

Oct 1, 2023

...

The content is so enlightening.



10



1 reply

[Reply](#)



Saurabh Chakravarty

Nov 9, 2023 (edited)

...

Typo : RHFL => RLHF



1



1 reply

[Reply](#)

Charleston Malkemus

Mar 7, 2024 (edited)

...

RHFL

Still not fixed RHFL >> RLHF



1 reply

[Reply](#)[See all responses](#)

More from Mastering LLM (Large Language Model)

MasteringLLM Blogs

11 Chunking Strategies for RAG—Simplified & Visualized

- 1. The Significance of Chunking
- 2. 11 Effective Chunking Strategies: Grasping the Process
- 3. Advantages and Disadvantages
- 4. Tips for Successful Execution
- 5. A Comparison of All 11 Chunking Strategies
- 6. How to Select the Right Strategy
- 7. Recommended Best Practices

www.masteringllm.com

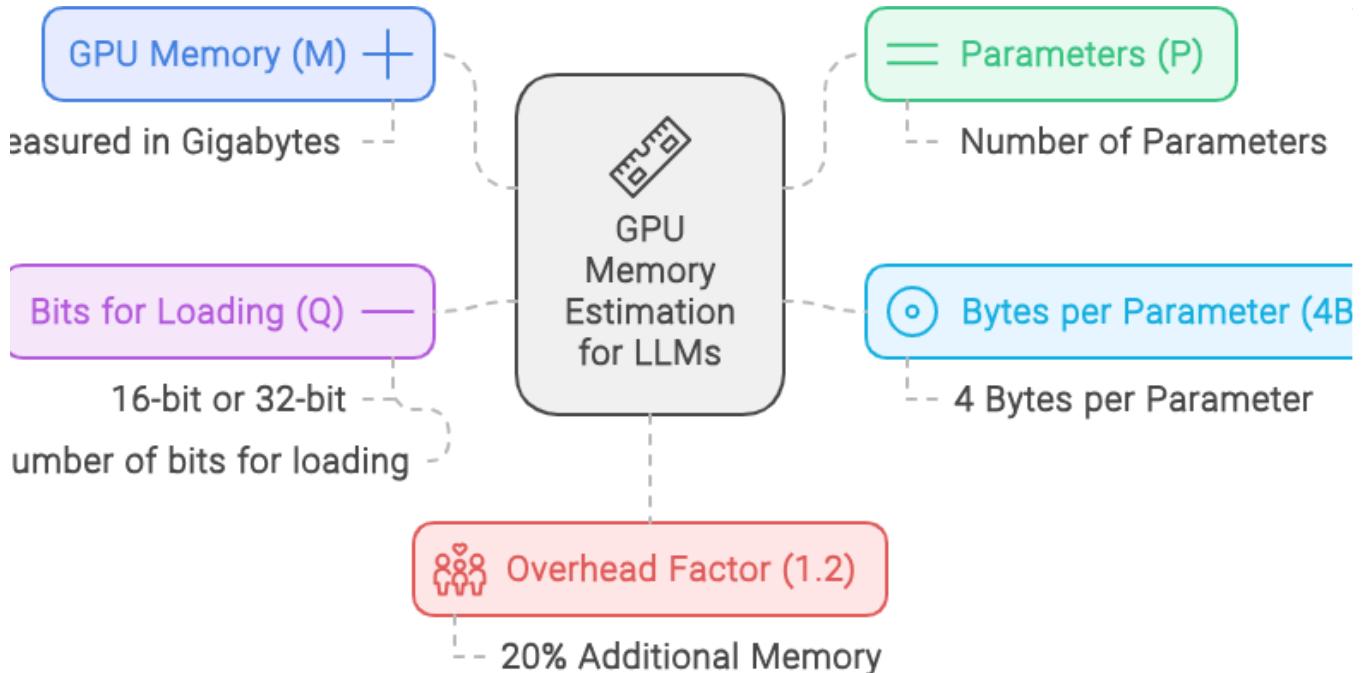
 Mastering LLM (Large Language Model)

11 Chunking Strategies for RAG—Simplified & Visualized

Retrieval-Augmented Generation (RAG) combines pre-trained language models with information retrieval systems to produce more accurate and...

Nov 2, 2024  169

...

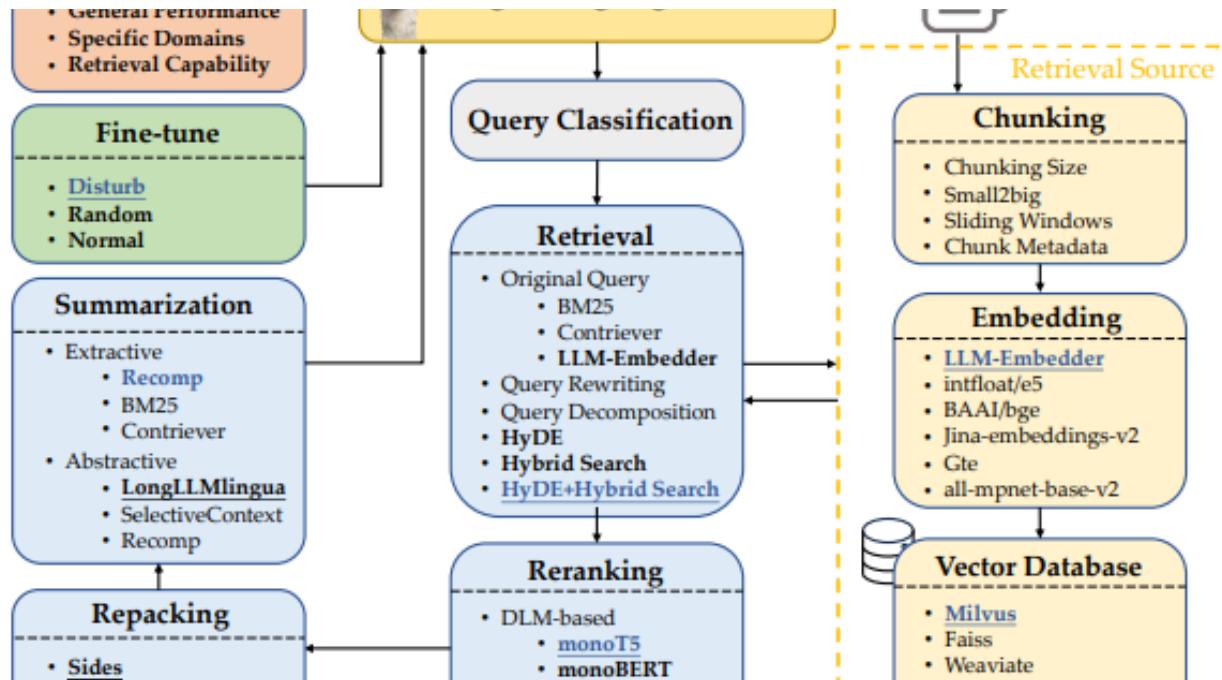

 Mastering LLM (Large Language Model)

How Much GPU Memory is Needed to Serve a Large Language Model (LLM)?

In nearly all LLM interviews, there's one question that consistently comes up: "How much GPU memory is needed to serve a Large Language Model..."

Aug 17, 2024  2K  20

...



Mastering LLM (Large Language Model)

Best Practices for RAG Pipeline

Over the past few years, RAG has matured and multiple studies have been done to understand patterns and behaviors which can result in low...

Sep 1, 2024

55



...

MasteringLLM Blogs

Mastering Caching Methods in Large Language Models (LLMs)

www.masteringllm.com

Mastering LLM (Large Language Model)

Mastering Caching Methods in Large Language Models (LLMs)

Large Language Models (LLMs) like OpenAI's GPT-4 have transformed natural language processing, enabling applications ranging from chatbots...

Sep 27, 2024

74



...

See all from Mastering LLM (Large Language Model)

Recommended from Medium

Always Free

24 GB RAM + 4 CPU + 200 GB



 @harendraverma2  @harendra21  @hendra21

 Harendra

How I Am Using a Lifetime 100% Free Server

Get a server with 24 GB RAM + 4 CPU + 200 GB Storage + Always Free

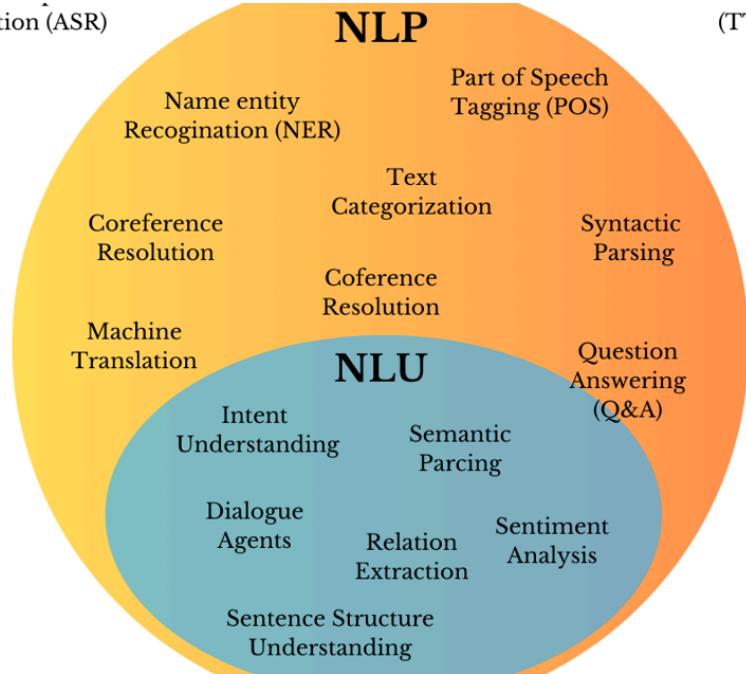
Oct 26, 2024

9.4K

170



...


 Vipra Singh

LLM Architectures Explained: NLP Fundamentals (Part 1)

Deep Dive into the architecture & building of real-world applications leveraging NLP Models starting from RNN to the Transformers.

Aug 15, 2024 2.6K 25



Lists



Predictive Modeling w/ Python

20 stories · 1856 saves



Natural Language Processing

1977 stories · 1619 saves



Practical Guides to Machine Learning

10 stories · 2225 saves



The New Chatbots: ChatGPT, Bard, and Beyond

12 stories · 563 saves



In Everyday AI by Manpreet Singh

Goodbye RAG? Gemini 2.0 Flash Have Just Killed It!

Alright!!!

Feb 10 2.8K 124



In Cubed by Michael Wood

The Insanity of Relying on Vector Embeddings: Why RAG Fails

In RAG, the goal is to locate the stored information that has the highest percentage of sameness to the provided query. Vector similarity...

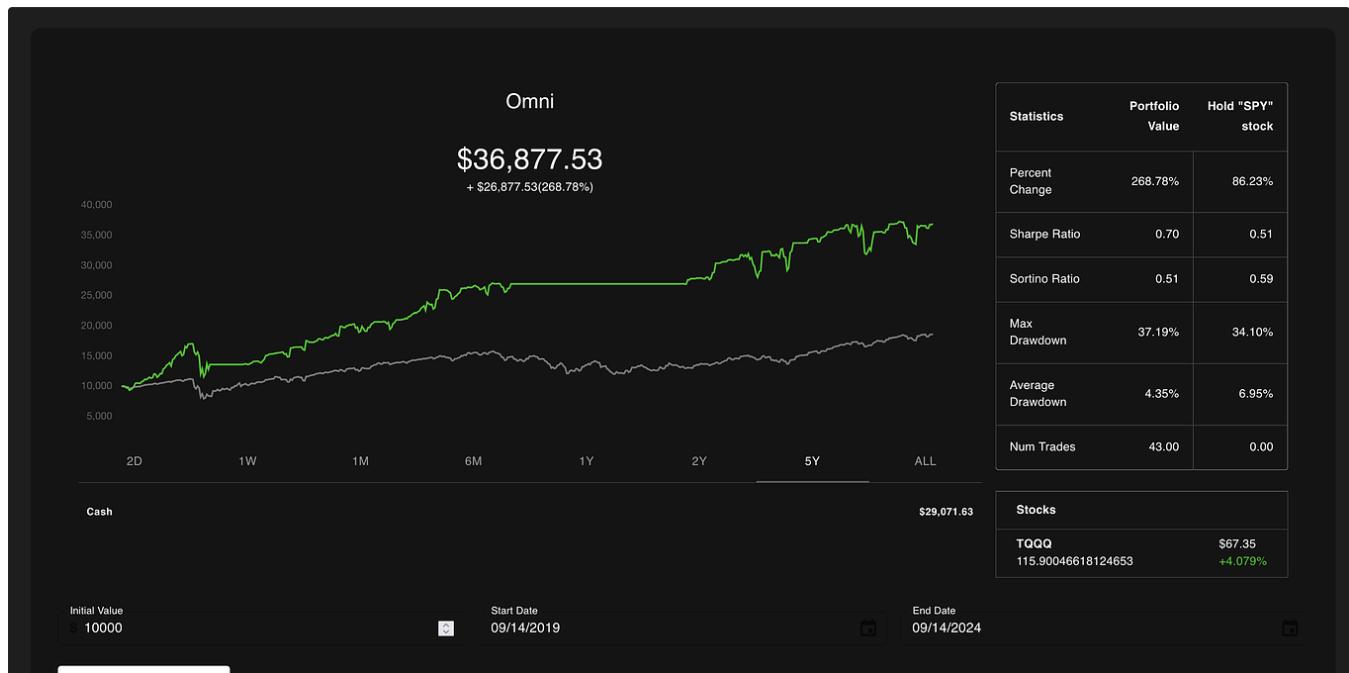
Nov 21, 2024

3.6K

71



...


 In DataDrivenInvestor by Austin Starks

I used OpenAI's o1 model to develop a trading strategy. It is DESTROYING the market

It literally took one try. I was shocked.

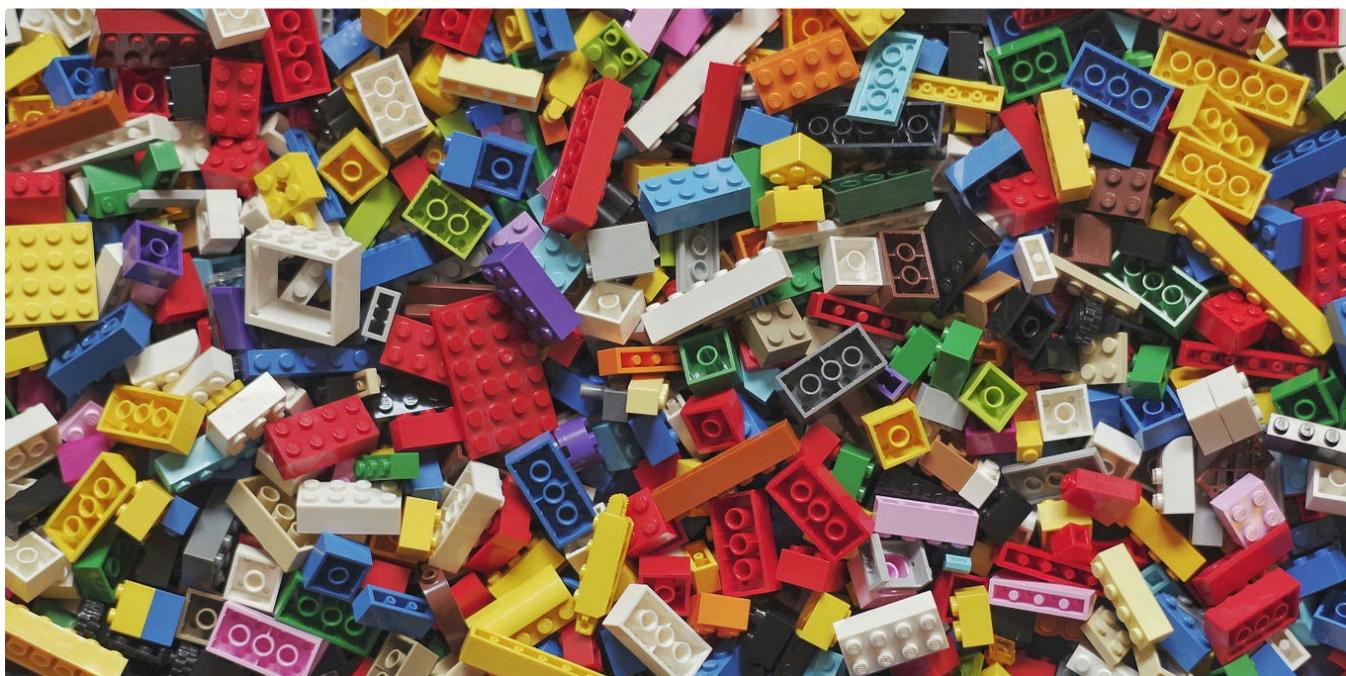
Sep 15, 2024

9K

241



...


 In Generative AI by Kenny Vaneetvelde 

Forget LangChain, CrewAI and AutoGen—Try This Framework and Never Look Back

In the rapidly evolving field of artificial intelligence, developers are inundated with frameworks and tools promising to simplify the...



Oct 21, 2024



2.3K



40



...

[See more recommendations](#)