✦ Member-only story

Building LLM Applications: Open-Source RAG (Part 7)

Vipra Singh · Follow

27 min read · Mar 16, 2024

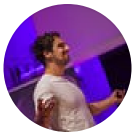
Learn Large Language Models (LLM) through the lens of a Retrieval Augmented Generation (RAG) Application.

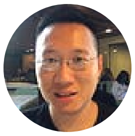
Posts in this Series

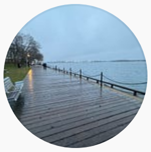
1. *Introduction*
2. *Data Preparation*
3. *Sentence Transformers*
4. *Vector Database*
5. *Search & Retrieval*
6. *LLM*
7. *Open-Source RAG (This Post)*
8. *Evaluation*

9. *Serving LLMs*
10. *Advanced RAG*

Table of Contents





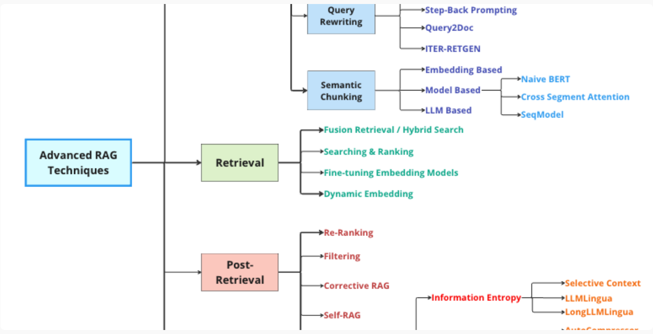


Written by Vipra Singh

1K Followers

Follow

More from Vipra Singh



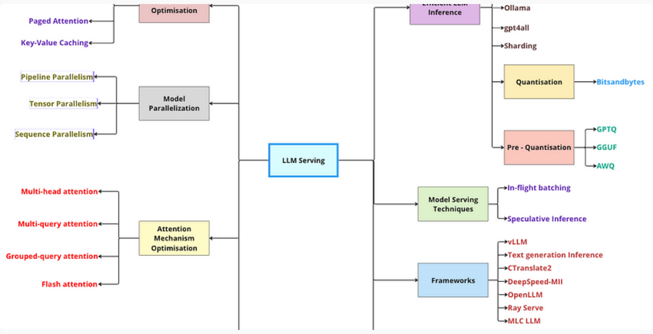
Vipra Singh

Building LLM Applications: Advanced RAG (Part 10)

Learn Large Language Models (LLM) through the lens of a Retrieval Augmented…

✦ · 48 min read · Apr 27, 2024

384 2



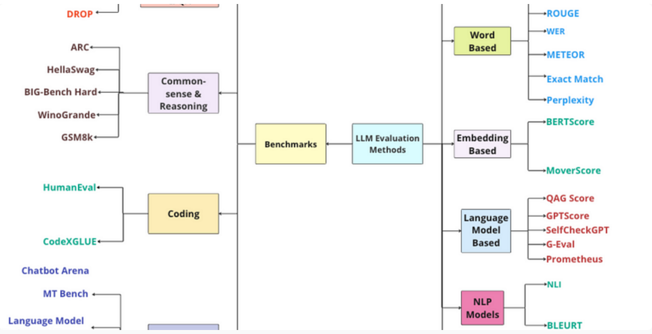
Vipra Singh

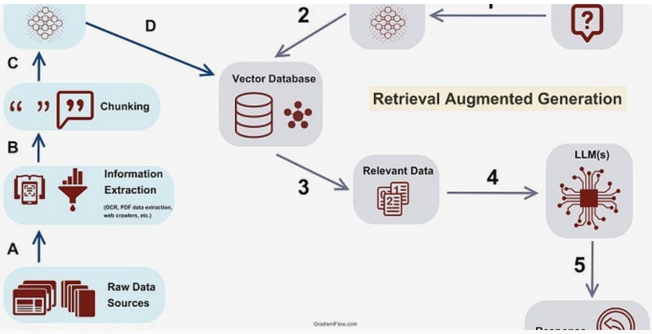
Building LLM Applications: Serving LLMs (Part 9)

Learn Large Language Models (LLM) through the lens of a Retrieval Augmented…

✦ · 50 min read · Apr 17, 2024

546 3





Vipra Singh

Building LLM Applications: Evaluation (Part 8)

Learn Large Language Models (LLM) through the lens of a Retrieval Augmented...

✦ · 48 min read · Apr 7, 2024

276 1

Vipra Singh

Building LLM Applications: Introduction (Part 1)

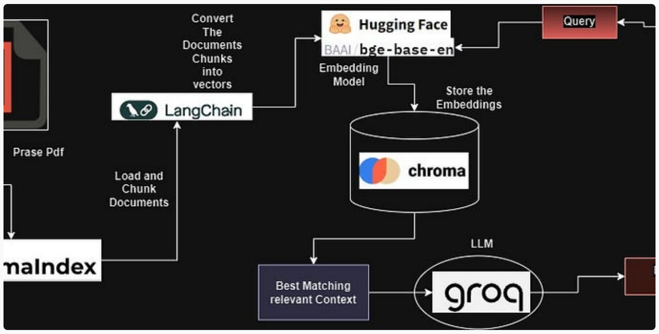
Learn Large Language Models (LLM) through the lens of a Retrieval Augmented...

5 min read · Jan 8, 2024

905 3

See all from Vipra Singh

Recommended from Medium



Plaban Nayak in The AI Forum

RAG on Complex PDF using LlamaParse, Langchain and Groq

Retrieval-Augmented Generation (RAG) is a new approach that leverages Large Languag…

13 min read · Apr 7, 2024

678 9

Paul Iusztin in Decoding ML

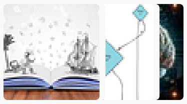
The 4 Advanced RAG Algorithms You Must Know to Implement

Implement from scratch 4 advanced RAG methods to optimize your retrieval and post-…

15 min read · May 4, 2024

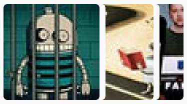
1.4K 10

Lists



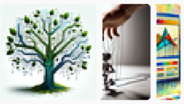
Natural Language Processing

1494 stories · 1011 saves



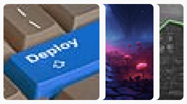
AI Regulation

6 stories · 473 saves



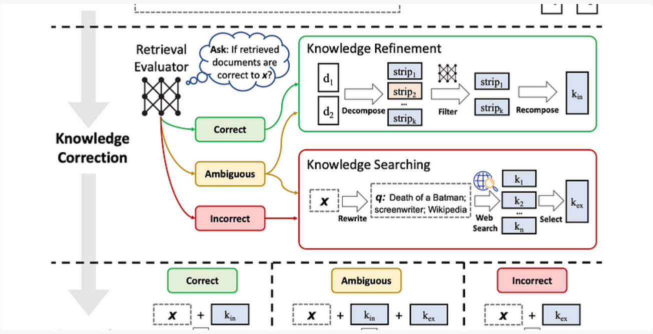
ChatGPT prompts

47 stories · 1642 saves



Predictive Modeling w/ Python

20 stories · 1254 saves

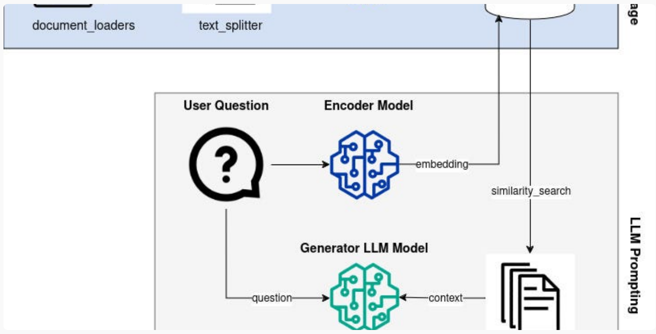


Barsha Rani Swain in GoPenAI

Advanced RAG: Corrective Retrieval Augmented Generation…

CRAG enhances the traditional RAG by introducing a retrieval evaluator to assess th…

10 min read · Apr 23, 2024



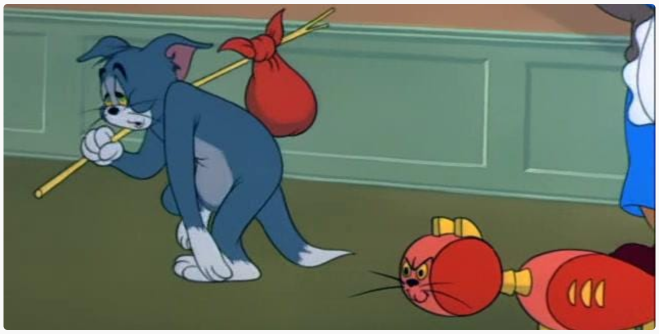
Dr. Leon Eversberg in Towards Data Science

How to Build a Local Open-Source LLM Chatbot With RAG

Talking to PDF documents with Google's Gemma-2b-it, LangChain, and Streamlit

✦ · 12 min read · Mar 31, 2024

320



Fareed Khan in Level Up Coding

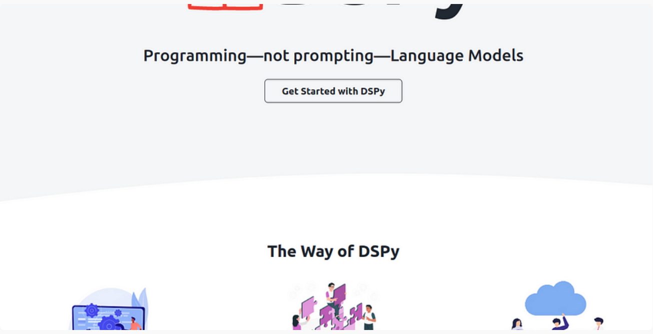
Building LLaMA 3 From Scratch with Python

Code Your Own Billion Parameter LLM

29 min read · May 28, 2024

1.1K 11

1.4K 8



Programming—not prompting—Language Models

Get Started with DSPy

The Way of DSPy

Vishal Rajput in AIGuys

Prompt Engineering Is Dead: DSPy Is New Paradigm For Prompting

DSPy Paradigm: Let’s program — not prompt — LLMs

✦ · 11 min read · May 29, 2024

1K 14

See more recommendations