

Team: Data Mavericks

Chengyi Cocoon CAO

Ansha Sondhi

Dhruv Choudhari

Nimagna Hazra

LLM Use-Case Real-Time Recommendation Engine

405 Data Management



INTRODUCTION TO OUR APP

- **LLM use-case search and recommendation engine** powered by **SingleStore's vector search** technology.
- **Hourly updates** with the latest data and user input from multiple sources
- Integrate **hybrid search** (full-text and vector) and **real-time analytics** to deliver instant and **user-specific** recommendations



APP FUNCTIONALITY & USE CASES

Functionality:

- LLM recommendations on basis of use case
- Recommendations ranked according to Hugging Face Leaderboard and key metrics like scores, likes and downloads
- Temporal relevance ensured through regular updating of usage and sentiment data from Github and Twitter
- Seamless user experience with free-text search functionality

Use Cases:

- Enabling writers, bloggers or marketers to select LLMs that are specialized for a particular style, theme or genre
- Assisting translators by recommending LLMs specialized in local languages, dialects and cultural nuances
- Supporting educators and curriculum designers by recommending LLMs suited for particular subjects like Calculus

Find the most appropriate LLM for your use case

I want an LLM that can help me support my customer queries related to product availability



01-ai/Yi-34B

🗨️ 0.7704 ❤️ 1129 ⬇️ 49255

An LLM is well-suited for your use case as it excels in generating human-like text by understanding context, tone, and style. Its advanced language capabilities make it an ideal tool for creating engaging and personalized content with minimal input.

Average	ARC	HellaSwag	MMLU	TruthfulQA	Winograde	GSM8K
69.42	64.59	85.69	76.35	56.23	83.03	50.64

[View details](#)

CausalLM/72B-preview

🗨️ 0.7670 ❤️ 71 ⬇️ 4381

The CausalLM/72B-preview model is well-suited for supporting customer queries related to product availability. Its advanced context understanding allows for accurate responses in a chat format, offering a seamless and interactive experience for users seeking information.

Average	ARC	HellaSwag	MMLU	TruthfulQA	Winograde	GSM8K
72.12	65.19	83.23	77.14	52.58	82.48	72.10

[View details](#)

mistralai/Mixtral-8x7B-v0.1

🗨️ 0.7646 ❤️ 1054 ⬇️ 113099

This LLM is well-suited for your use case because of its robust natural language understanding capabilities, allowing it to accurately interpret and generate human-like text. Its advanced algorithms can handle complex language tasks with high accuracy, making it an ideal choice...

Average	ARC	HellaSwag	MMLU	TruthfulQA	Winograde	GSM8K
68.42	66.04	86.49	71.82	46.78	81.93	57.47

[View details](#)

TECHNICAL ARCHITECTURE



Data Sources & Processing:

- Multiple API Integrations
 - Hugging Face Leaderboard: LLM models
 - X: user feedback
 - Github: repos, use cases
- SingleStore Ecosystem

Database & Search Capabilities:

- SingleStore Database
- Search interface

Application Deployment:

- Front-End Application
 - Elegance SDK and Vercel

CHALLENGES & SOLUTIONS

CHALLENGES	SOLUTIONS
Issues with ingesting GitHub Data through APIs	Created a a personal access token and updated the required permissions
Issues with ingesting X Data through user APIs	Would require payment to use the API to access X data
Inability to populate OpenAI embeddings in dataset	Resolved issues with API security access. Created new account to get additional free credits
Hosting complications due to multiple cross-platform integrations	Implementing unified tech stacks that allow seamless deployment

SINGLESTORE'S VECTOR SEARCH

The platform has a efficient indexing mechanism which streamlines the implementation process. Unlike traditional vector databases, SingleStore stores vector data in relational tables alongside other types of data.

ADVANTAGES

- **Nearest neighbor search:** Since SingleStore supports joins, you can do set-based nearest-neighbor search in place of doing multiple queries to find desired results.
- **Hybrid Search Capabilities:** descriptive properties is easy to use in SingleStore, because all the query capabilities of SQL are available.
- **Semantic Search:** Semantic search capabilities allow you to build applications based on LLMs that are capable of finding text that matches the meaning of your query, not just the words it contains

DISADVANTAGES

- **Configuration Complexity:** Adjusting how things are grouped together and deciding how similar they should be, needs careful testing which makes the setup a bit more complicated.
- **Learning Curve:** Learning how to use the new way of searching with vectors can be difficult, but the instructions from SingleStore helps the process.

**THANK
YOU**
