

I use Cosine distance, however, I found a strange problem, when I use '印度' as search sentence, the return result show a opposite situation:

1. the return result which has larger distance value contain the '印度' word.
2. the return result which has small distance value **doen not** contain the '印度' word.

This is so different from the above discussion.

<think>

本題描述使用 pymilvus 進行 cosine distance 搜尋時，發現搜尋「印度」時，距離值較大的結果反而包含「印度」，而距離值較小的結果卻沒有「印度」，這與 cosine distance 越小越相似的預期相反。首先需確認 cosine distance 在 Milvus 的定義與實作，根據官方文件與 GitHub Issue，Milvus 對 cosine distance 的回傳值有異常現象：有用戶回報距離值出現「-1」代表完全相同，0 代表不相關，1 代表完全相反，這與一般 cosine similarity (1=完全相同，-1=完全相反) 定義相反，且距離值排序也可能顛倒^[1]。

此外，Milvus 官方文件指出 cosine similarity 可由 1 減去 cosine similarity 得到 cosine distance^[2]，但實際上在某些版本或情境下，Milvus 可能直接回傳 cosine similarity 或其負值^[1]。

再者，若 embedding 本身對於非英文文本（如「印度」）表現不佳，也可能導致語意無法正確反映在向量空間，進而影響搜尋結果^[3]。

綜合來看，這個問題可能同時涉及 Milvus 實作 bug、distance 定義混淆，以及 embedding 模型對多語言文本的支援度。

最終需明確告知用戶，這屬於 Milvus cosine distance 回傳異常的已知問題，建議關注官方修正進度，或考慮 workaround，例如自行後處理 distance 值排序。

</think>

根據你描述的現象，這其實是 Milvus 在 cosine distance 計算上已知的異常問題。根據 2024 年 5 月的 GitHub 回報，當你設定 metric_type 為 COSINE 時，Milvus 回傳的 distance 值與一般認知的 cosine distance 定義相反：

- 完全相同的向量會得到 -1
- 完全不相關得到 0
- 完全相反得到 1

也就是說，目前 Milvus 回傳的 distance 其實是「cosine similarity 的負值」或是直接回傳 similarity 而非 distance，導致排序與預期顛倒^[1]。

這會造成你遇到的情況：「距離值較大」的結果反而更相關，因為在這個實作下，越小的 distance 其實代表越不相似，越大的 distance 反而是越相似。這與 cosine distance 應該「越小越相似」的定義是相反的^{[1] [2]}。

解決建議

- **暫時解法**：你可以在應用層將 distance 乘以 -1 或重新排序，讓結果符合「越小越相似」的直覺。
- **關注官方修正**：這個問題已被官方回報，建議持續關注 Milvus 的 issue 更新^[1]。
- **檢查 embedding 模型**：如果你用的 embedding 模型對非英文文本（如「印度」）支援不佳，也可能導致語意不準確，建議確認 embedding 輸出的品質^[3]。

補充說明

Milvus 官方文件也明確指出，cosine similarity 介於 $[-1, 1]$ ，而 cosine distance 應為 $1 - \text{cosine similarity}$ ^[2]。但目前實作與文件描述不符，這正是你遇到問題的根本原因。

> 「It looks that the distance value returned from the search on a collection of vectors with a cosine metric similarity is wrong. It's the opposite of what it should be...」^[1]

結論

你遇到的現象並非你的應用邏輯錯誤，而是 Milvus cosine distance 回傳異常的已知問題。建議暫時在應用層自行修正排序，並密切關注官方後續修正。

✱

1. <https://github.com/milvus-io/milvus/issues/33072>
2. <https://milvus.io/docs/metric.md>
3. <https://stackoverflow.com/questions/77179433/cosine-similarity-search-with-non-english-text-giving-wrong-results>