

RAG 系統評估報告

文本切塊方法參數與結果

一、固定大小切塊和滑動視窗切塊的參數設定

| 切塊方法 | 參數名稱 | 參數值 |
|--------|---------------|---------------------|
| 固定大小切塊 | chunk_size | 300 字元 |
| | chunk_overlap | 10 字元 |
| | separator | 。 |
| 滑動視窗切塊 | chunk_size | 350 tokens |
| | chunk_overlap | 15 tokens |
| | separators | ["。", "\n\n", "\n"] |

二、效果最好的切塊方法

滑動視窗切塊方法 (Sliding Window) 效果最好，平均分數為 0.7079。