

какие-то общие слова щас скажу на сорева по рагу включайте голову, не собирайте франкенштейна без надобности. сперва, посмотрите на документ который вам дали, оцените его структуру. прикиньте какой чанкер и тд будет лучше себя вести на нем, можно ли родительские тексты норм получить, можно ли мету интересную собрать потом соберите бейз и зашлите посылку потом посмотрите на вывод каждого блока (дима это сделал), посмотрите что первая стадия ретривала дает, что вторая, что реранкер, какой контекст приходит в итоге модели. и в итоге вы локализуете ошибку и бейте улучшениями в неё. если вы видите что вроде бы все дали но генерится фигня, улучшайте промпты

Соревнования <https://www.kaggle.com/competitions/casml-generative-ai-hackathon>
<https://www.kaggle.com/competitions/lhp-rag-hackathon> - чисто ретривал
<https://www.kaggle.com/competitions/financial-rag-implementation> - чисто ретривал
<https://www.kaggle.com/competitions/icaif-24-finance-rag-challenge/leaderboard>
<https://www.kaggle.com/competitions/bigquery-ai-hackathon/writeups/causal-rag> - почитать
<https://www.kaggle.com/competitions/rag-challenge-madagascar-education-q-a-public>
<https://habr.com/ru/articles/893356/> - почитать

https://github.com/NirDiamant/RAG_Techniques