<https://www.kaggle.com/datasets/mikhailklemin/kinopoisks-movies-reviews?resource=download>

**Kinopoisk's movies reviews**

В вопросе выбора датасета и инструментов для выполнения задания

ограничений нет

Ожидается, что на вход программа принимает такие параметры, как поисковый запрос,

количество результатов для выдачи и метод поиска.

На выходе - топ документов в порядке убывания релевантности, время исполнения программы,

а также интерпретируемые метрики ранжирования, отражающие общее качество модели на

датасете и уверенность в возвращаемых результатах.

В методах поиска должны быть следующие варианты:

1. полнотекстовый

2. векторный

3. гибридный

Для векторного поиска обосновать выбор той или иной модели. Для гибридного - описать

использованный алгоритм объединения результатов. Раскрыть детали предобработки данных,

если таковая проводилась.

Интерфейс: у программы должен быть хорошо задокументированный user-friendly API.

Деплой: Контейнеризовать проект (собрать docker-образ приложения), написать инструкции для

запуска.

Оценка: Протестировать реализованные методы поиска и сформировать сопоставительный

отчёт/дашборд по качеству поисковика на разных движках (т.е. провалидировать и

интерпретировать решения).