

## ASSIGNMENT FOR BACK-END POSITION\_DISCOVERY TEAM

Xây dựng một small search engine bởi hai phương pháp cơ bản, dữ liệu đầu vào gồm các tên của sản phẩm trong file `product\_names.txt` (gồm 50000 tên của sản phẩm).

- (Keyword search) Sử dụng reverse index để tìm kiếm
- (Fulltext search) Sử dụng thuật toán BM25 để tìm kiếm
- So sánh hai mô hình đã viết ở a. và b. chạy kết quả cho 100 query trong file query.txt

Bạn nên viết các hàm để preprocessing document và query, và hỗ trợ cho việc tìm kiếm không dấu.

