

Thực hành đánh giá truy hồi thông tin

Nguyễn Mạnh Hiễn
hiennm@tlu.edu.vn

Đề bài

Lập trình Python để tính độ đo NDCG theo các bước sau:

1. Nhập vào danh sách (mã) văn bản và độ phù hợp kèm theo. Gợi ý: Có thể dùng kiểu danh sách (list) trong Python, mỗi phần tử của danh sách là một bộ (tuple) gồm mã văn bản và độ phù hợp.
2. Tính thu hoạch tích lũy CG tại các hạng khác nhau.
3. Tính thu hoạch tích lũy có khấu trừ DCG tại các hạng khác nhau.
4. Tính NDCG tại các hạng khác nhau:
 - 4.1. Tạo danh sách lý tưởng rồi lặp lại các tính toán như trong các bước 2 và 3 để thu được IDCG (DCG lý tưởng). Gợi ý: Nên viết các bước 2 và 3 thành hàm để tránh lặp code.
 - 4.2. Tính NDCG bằng cách chia DCG cho IDCG.

Khi chạy thử, hãy nhập dữ liệu như trong slide 19 của bài giảng lý thuyết.