



IT4853

Tìm kiếm và trình diễn thông tin

Giới thiệu môn học



Thông tin giảng viên

- TS. Nguyễn Kiên Hiếu
- BM Hệ thống thông tin, Viện CNTT-TT
- hieunk@soict.hust.edu.vn
- <https://users.soict.hust.edu.vn/hieunk/>



Đánh giá môn học

- Điểm quá trình (30%): BTL (theo nhóm)
- Điểm cuối kỳ (70%): Tự luận



Tài liệu tham khảo

- [1] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- [2] W. Bruce Croft, Donald Metzler and Trevor Strohman (2010). *Search Engines: Information retrieval in practice*. Addison-Wesley.



Đề cương môn học

- Bài 1 – Các khái niệm cơ bản
- Bài 2 – Thực hiện truy vấn Boolean trên chỉ mục ngược
- Bài 3 – Xử lý từ truy vấn
- Bài 4 – Mô hình không gian véc-tơ
- Bài 5 - Mô hình tìm kiếm xác suất
- Bài 6 - Mô hình ngôn ngữ
- Bài 7, 8 - Đánh giá kết quả tìm kiếm



Đề cương môn học

- Bài 9 - Nén chỉ mục ngược
- Bài 10 - Các phương pháp xây dựng chỉ mục ngược
- Bài 11, 12 - Phân lớp văn bản
- Bài 13, 14 - Phân cụm văn bản
- Bài 15 - Vấn đề tìm kiếm trên web
- Bài 16 - Phát hiện trùng lặp gần
- Bài 17 - Quảng cáo và spam



Đề cương môn học

- Bài 18 - Thu thập dữ liệu
- Bài 19 - Phân tích liên kết – Pagerank
- Bài 20 - Phân tích liên kết - HITS