

TUGAS #1

Teks dan Web Mining

Semester Genap Tahun Ajaran 2013/2014

Tujuan

Agar mahasiswa dapat memahami metode evaluasi dalam sistem information retrieval dengan menggunakan alat ukur bakuyaitu precision, recall, dan F1 Measure.

Deadline

Laporan tugas dikumpulkan melalui e-Class dengan format file PDF paling lambat tanggal 28 Maret 2014 pukul 23:59 WIB. Nama file mengikuti format <4nimakhir>_<4nimakhir>_<4nimakhir>_<4nimakhir>.pdf. **Jika nama file tidak sesuai format, maka akan ditolak.**

Deskripsi Pekerjaan

Buatlah sebuah query string yang akan Anda berikan kepada mesin pencari, yaitu Google, Bing, dan Yahoo!. Query string tersebut harus sesuai dengan kebutuhan informasi dari topik studi kasus yang telah ditentukan untuk masing-masing kelompok.

Misal, query string yang akan diberikan adalah "Kota Terpadat di Indonesia". Maka query string tersebut akan dikenakan ke Google dan juga Bing dan atau Yahoo!.

Penentuan mesin pencari tiap kelompok akan ditentukan dengan aturan sebagai berikut:

- Nomor kelompok mod 3 = 0 akan menggunakan mesin pencari Google dan Bing
- Nomor kelompok mod 3 = 1 akan menggunakan mesin pencari Google dan Yahoo!
- Nomor kelompok mod 3 = 2 akan menggunakan mesin pencari Bing dan Yahoo!

Untuk masing-masing hasil yang dikembalikan oleh mesin pencari, ambil 30 dokumen pertama, sehingga total maksimal adalah 60 dokumen. Dari 30 dokumen pertama untuk masing-masing mesin pencari, tentukan dokumen mana saja yang relevan dengan topik yang Anda butuhkan. Buatlah daftar dokumen-dokumen yang relevan tersebut sesuai urutan tampilan dari mesin pencari. (Pemilihan dokumen yang relevan dapat menggunakan mekanisme pooling dari anggota kelompok).

Dari hasil tersebut, buatlah tabel precision dan recall dengan menggunakan tabel berikut ini:

	<Mesin Pencari#1>			<Mesin Pencari#2>		
	P	R	F1	P	R	F1
Top 5						
Top 10						
Top 15						
Top 20						
Top 25						
Top 30						

Gambarkan grafik precision, recall, dan f1-measure berdasar temuan tersebut!