

**PERBANDINGAN SISTEM TEMU KEMBALI MODEL
RUANG VEKTOR DENGAN SISTEM TEMU
KEMBALI OKAPI BM25
PADA PERANKINGAN DOKUMEN TEXT**

Tugas Akhir



Oleh :

Benyamin Kurniadi G

22074268

**Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana Yogyakarta
Tahun 2011**

**PERBANDINGAN SISTEM TEMU KEMBALI MODEL RUANG
VEKTOR DENGAN SISTEM TEMU KEMBALI OKAPI BM25
PADA PERANKINGAN DOKUMEN TEXT**

Tugas Akhir



Diajukan kepada Fakultas Teknologi Informasi Teknik Informatika

Universitas Kristen Duta Wacana Sebagai salah satu syarat

dalam memperoleh gelar Sarjana Komputer



Oleh :

Benyamin Kurniadi G

22074268

Program Studi Teknik Informatika Fakultas Teknologi Informasi

Universitas Kristen Duta Wacana Yogyakarta

Tahun 2011

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa tugas akhir dengan judul :

Perbandingan Sistem Temu Kembali Model Ruang Vektor dengan Sistem Temu Kembali OKAPI BM25 pada Perangkingan Dokumen Text

Yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana komputer pada pendidikan sarjana Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 27 April 2011



Benyamin Kurniadi Gunawan

22074268



HALAMAN PERSETUJUAN

Judul : Perbandingan Sistem Temu Kembali Model Ruang Vektor
dengan Sistem Temu Kembali OKAPI BM25 pada
Perangkingan Dokumen Text

Nama : Benyamin Kurniadi Gunawan

NIM : 22074268

Mata Kuliah : Tugas Akhir

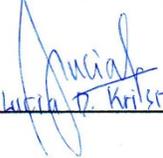
Kode : TI2116

Semester : Gasal

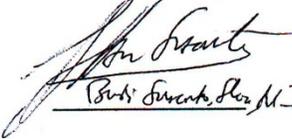
Tahun Akademik : 2010/2011

Telah diperiksa dan disetujui
Di Yogyakarta,
Pada Tanggal 27 April 2011

Dosen Pembimbing I


Nuria P. Kristawati, MA

Dosen Pembimbing II


Budi Suwanto, S.S., M.T.

HALAMAN PENGESAHAN

SKRIPSI

Perbandingan Sistem Temu Kembali Model Ruang Vektor

dengan Sistem Temu Kembali OKAPI BM25 pada

Perangkingan Dokumen Text

Oleh : Benyamin Kurniadi Gunawan

Dipertahankan di depan dewan Penguji Tugas Akhir/Skripsi Program Studi Teknik Informatika
Fakultas Teknologi Informasi Universitas Kristen Duta Wacana – Yogyakarta dan dinyatakan diterima
memenuhi salah satu syarat memperoleh gelar Sarjana Komputer

Pada Tanggal
9 Mei 2011

Yogyakarta, 18 Mei 2011

Mengesahkan,

Dewan Penguji :

1. Lucia Dwi Krisnawati, S.S., M.A.
2. Budi Susanto, S.Kom, M.T.
3. Willy Sudiarto R, S.Kom., M.Cs.
4. Joko Purwadi, S.Kom., M.Kom.



Dekan

Drs. Wimmie Handiwidjojo, MIT.

Ketua Program Studi

Nugroho Agus H,S.Si.,M.Si.

UCAPAN TERIMA KASIH

Puji syukur kepada Tuhan atas segala berkat, bimbingan, dan perlindungan-Nya selama mengerjakan Tugas Akhir ini sehingga dapat diselesaikan dengan baik dan tepat pada waktunya.

Penulisan laporan ini merupakan kelengkapan dan pemenuhan dari salah satu syarat dalam memperoleh gelar Sarjana Komputer. Selain itu bertujuan melatih siswa untuk dapat menghasilkan suatu karya yang dapat dipertanggung jawabkan secara ilmiah, sehingga dapat bermanfaat bagi penggunanya.

Pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada pihak-pihak yang telah banyak memberikan bimbingan, bantuan dan dukungan, antara lain:

1. **Ibu Lucia Dwi Krinawati, M.A.** selaku dosen pembimbing I yang telah memberikan bimbingannya dengan sabar dan baik kepada penulis.
2. **Bpk Budi Susansto, S.Kom, M.T.** selaku dosen pembimbing II atas petunjuk dan masukan yang diberikan selama pengerjaan tugas ini sejak awal hingga akhir.
3. Ayah dan Ibu, serta keluarga yang selalu memberi dukungan, do'a dan kasih sayang bagi penulis sehingga tugas ini dapat terselesaikan dengan baik.
4. Rekan-rekan dan pihak-pihak yang tidak dapat penulis sebutkan satu persatu yang secara langsung maupun tidak langsung yang telah mendukung penyelesaian tugas ini. Terima kasih atas dukungan dan do'anya.
5. Yang terakhir dan yang terpenting, kepada Tuhan Yesus Kristus, yang memberikan seluruh kekuatan, waktu, dan kesehatan selama penulis mengerjakan kerja praktek ini.

Dengan tersusunnya laporan ini, penulis berharap dapat memberikan informasi dan wawasan kepada pembaca

Penulis menyadari bahwa laporan yang penulis susun ini masih banyak kekurangannya. Oleh karena itu, penulis mohon saran dan kritik dari pembaca guna kesempurnaan tulisan ini.

Yogyakarta, Mei 2011

Penulis

© UKDW

INTISARI

PERBANDINGAN SISTEM TEMU KEMBALI MODEL RUANG VEKTOR DENGAN SISTEM TEMU KEMBALI OKAPI BM25 PADA PERANKINGAN DOKUMEN TEXT

Sekarang ini sistem IR (*Information Retrieval*) telah menjadi solusi dalam melakukan pencarian dokumen dengan jumlah yang besar Metode yang digunakan untuk membangun sistem IR sendiri semakin bervariasi. Salah satunya adalah metode OKAPI BM25 dan metode Model Ruang Vektor (*Vector Space Model*, kedua metode tersebut dikenal dapat menghasilkan tingkat presisi yang tinggi dari hasil pencarian yang dilakukan. Penulis merasa perlu membandingkan kedua metode tersebut untuk mengetahui metode mana yang dapat menghasilkan rata-rata presisi yang lebih tinggi pada proses perankingan dokumen text.

Untuk menemukan solusi dari masalah yang diangkat pada penelitian penulis mengimplentasikan sistem temu kembali OKAPI BM25 dan sistem temu kembali Model Ruang Vektor. Pada kedua sistem yang dibangun dilakukan proses eliminasi *stopword* sebelum dilakukan proses indexing. Index database yang dihasilkan dari proses indexing digunakan pada proses pencarian dokumen relevan. Oleh masing-masing sistem hasil pencarian kemudian diurutkan berdasarkan nilai kesamaan untuk mendapatkan ranking dokumen relevan, dari hasil pencarian kedua sistem tersebut dihitung besarnya nilai presisi dan *recall* untuk digunakan pada proses evaluasi.

Dari hasil analisis dan evaluasi terhadap kedua sistem temu kembali dengan menggunakan delapan *query* dan 105 dokumen text yang disiapkan penulis, ternyata metode OKAPI BM25 memiliki nilai rata-rata presisi yang lebih tinggi daripada metode Model Ruang Vektor.

Keyword : *information retrieval, OKAPI BM25, Model Ruang Vektor*

DAFTAR ISI

HALAMAN JUDUL	
PERNYATAAN KEASLIAN SKRIPSI	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
UCAPAN TERIMA KASIH	iv
INTISARI	vi
DAFTAR ISI	vii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.3.1 Format Input	2
1.3.2 Fromat Output	3
1.4 Hipotesis	3
1.5 Tujuan	3
1.6 Metode Penelitian	4
1.7 Sistematika Penulisan	4
BAB 2 TINJAUAN PUSTAKA	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori	8

2.2.1 Indexing	8
2.2.2 <i>Stopwords</i>	9
2.2.3 TF-IDF	9
2.2.4 Metode OKAPI BM25.....	11
2.2.5 Metode Model Ruang Vektor.....	13
2.2.5 <i>Information Retrieval</i>	15
BAB 3 ANALISIS DAN PERANCANGAN SISTEM	17
3.1 Kebutuhan <i>Hardware</i> dan <i>Software</i>	17
3.2 Spesifikasi Sistem	18
3.3 Arsitektur Sistem	19
3.3.1 Sistem Temu Kembali Model Ruang Vektor	19
3.3.2 Sistem Temu Kembali OKAPI BM25	20
3.4 <i>Use Case</i>	21
3.5 Algoritma dan <i>Flowchart</i>	22
3.5.1 Tambah Dokumen	22
3.5.2 Tambah <i>Stopwords</i>	23
3.5.3 Perangkingan Model Ruang Vektor	25
3.5.4 Perangkingan OKAPI BM25	28
3.6 Kamus Data	30
3.6.1 Tabel File	30
3.6.2 Tabel Token	30
3.6.3 Tabel File_has_token	31
3.6.4 Tabel Stopword	32
3.6.5 Tabel Admin	32
3.7 Diagram Skema	33
3.8 Rancangan Antarmuka Sistem	14
3.8.1 Rancangan Halaman Utama	34
3.8.2 Rancangan Antarmuka Sistem Temu Kembali Model	

Ruang Vektor	35
3.8.3 Rancangan Antarmuka Sistem Temu Kembali OKAPI BM25	37
3.8.1 Rancangan Halaman Tambah Dokumen	38
3.8.1 Rancangan Halaman Tambah Stopwords	39
3.9 Rancangan Pengujian Sistem	40
3.9.1 Sistem Temu Kembali Model Ruang Vektor	40
3.9.2 Sistem Temu Kembali OKAPI BM25	41
BAB 4 IMPLEMENTASI DAN ANALISIS SISTEM	42
4.1 Implementasi Sistem	42
4.1.1 Halaman Utama	42
4.1.2 Halaman Sistem Temu Kembali Model Ruang Vektor	43
4.1.3 Halaman Sistem Temu Kembali OKAPI BM25	46
4.1.4 Halaman Login	48
4.1.5 Halaman Tambah Dokumen	49
4.1.5.1 <i>Pseudo Code</i> Proses Tambah Dokumen	51
4.1.5.2 <i>Pseudo Code</i> Proses Indexing	52
4.1.6 Halaman Tambah Stopwords	53
4.2 Analisis Sistem	54
4.2.1 Pengujian Sistem Temu Kembali Model Ruang Vektor	55
4.2.1.1 Pengujian Pertama	55
4.2.1.2 Pengujian Kedua	56
4.2.1.3 Pengujian Ketiga	57
4.2.1.4 Pengujian Keempat	58
4.2.1.5 Pengujian Kelima	59
4.2.1.6 Pengujian Keenam	60
4.2.1.7 Pengujian Ketujuh	61
4.2.1.8 Pengujian Kedelapan	62

4.2.1.9 Contoh Grafik <i>Interpolated Precision</i>	63
4.2.1.10 Rata-rata Hasil Pengujian	65
4.2.2 Pengujian Sistem Temu Kembali OKAPI BM25	67
4.2.2.1 Pengujian Pertama	67
4.2.2.2 Pengujian Kedua	68
4.2.2.3 Pengujian Ketiga	69
4.2.2.4 Pengujian Keempat	70
4.2.2.5 Pengujian Kelima	71
4.2.2.6 Pengujian Keenam	72
4.2.2.7 Pengujian Ketujuh	73
4.2.2.8 Pengujian Kedelapan	74
4.2.2.9 Contoh Grafik <i>Interpolated Precision</i>	75
4.2.2.10 Rata-rata Hasil Pengujian	77
4.3 Hasil Perbandingan	78
BAB 5 KESIMPULAN DAN SARAN	80
5.1 Kesimpulan	80
5.2 Saran	81



DAFTAR TABEL

Tabel 2.1	Pembobotan TF-IFD	4
Tabel 3.1	Tabel File	30
Tabel 3.2	Tabel Token	31
Tabel 3.3	Tabel File_has_token	31
Tabel 3.4	Tabel Stopwords	32
Tabel 3.5	Tabel Admin	32
Tabel 4.1	Daftar Dokumen Relevan	54
Tabel 4.2	Hasil Pengujian Pertama	55
Tabel 4.3	<i>Eleven Points Interpolated Precision P1</i>	55
Tabel 4.4	Hasil Pengujian Kedua	56
Tabel 4.5	<i>Eleven Points Interpolated Precision P2</i>	56
Tabel 4.6	Hasil Pengujian Ketiga	57
Tabel 4.7	<i>Eleven Points Interpolated Precision P3</i>	57
Tabel 4.8	Hasil Pengujian Keempat	58
Tabel 4.9	<i>Eleven Points Interpolated Precision P4</i>	58
Tabel 4.10	Hasil Pengujian Kelima	59
Tabel 4.11	<i>Eleven Points Interpolated Precision P5</i>	59
Tabel 4.12	Hasil Pengujian Keenam	60
Tabel 4.13	<i>Eleven Points Interpolated Precision P6</i>	60
Tabel 4.14	Hasil Pengujian Ketujuh	61
Tabel 4.15	<i>Eleven Points Interpolated Precision P7</i>	61
Tabel 4.16	Hasil Pengujian Kedelapan	62
Tabel 4.17	<i>Eleven Points Interpolated Precision P8</i>	62
Tabel 4.18	Rata-Rata Hasil Pengujian Sistem Temu Kembali Model Ruang Vektor	66

Tabel 4.19	Hasil Pengujian Pertama	67
Tabel 4.20	<i>Eleven Points Interpolated Precision P1</i>	67
Tabel 4.21	Hasil Pengujian Kedua	68
Tabel 4.22	<i>Eleven Points Interpolated Precision P2</i>	68
Tabel 4.23	Hasil Pengujian Ketiga	69
Tabel 4.24	<i>Eleven Points Interpolated Precision P3</i>	69
Tabel 4.25	Hasil Pengujian Keempat	70
Tabel 4.26	<i>Eleven Points Interpolated Precision P4</i>	70
Tabel 4.27	Hasil Pengujian Kelima	71
Tabel 4.28	<i>Eleven Points Interpolated Precision P5</i>	71
Tabel 4.29	Hasil Pengujian Keenam	72
Tabel 4.30	<i>Eleven Points Interpolated Precision P6</i>	72
Tabel 4.31	Hasil Pengujian Ketujuh	73
Tabel 4.32	<i>Eleven Points Interpolated Precision P7</i>	73
Tabel 4.33	Hasil Pengujian Kedelapan	74
Tabel 4.34	<i>Eleven Points Interpolated Precision P8</i>	74
Tabel 4.35	Rata-Rata Hasil Pengujian Sistem Temu Kembali OKAPI BM25	78



DAFTAR GAMBAR

Gambar 3.1	Arsitektur Sistem Temu Kembali Model Ruang Vektor ..	19
Gambar 3.2	Arsitektur Sistem Temu Kembali OKAPI BM25	20
Gambar 3.3	Diagram Use Case	21
Gambar 3.4	<i>Flowchart</i> Tambah Dokumen	23
Gambar 3.5	<i>Flowchart</i> Tambah <i>Stopwords</i>	24
Gambar 3.6	<i>Flowchart</i> Proses Indexing	26
Gambar 3.7	<i>Flowchart</i> Proses Perangkingan	27
Gambar 3.8	<i>Flowchart</i> Proses Perangkingan OKAPI BM 25	29
Gambar 3.9	Diagram Skema	33
Gambar 3.10	Rancangan Halaman Utama	34
Gambar 3.11	Rancangan Halaman Sistem Temu Kembali Model Ruang Vektor	35
Gambar 3.12	Rancangan Halaman Login	36
Gambar 3.13	Rancangan Halaman Sistem Temu Kembali OKAPI BM25	37
Gambar 3.14	Rancangan Halaman Tambah Dokumen	38
Gambar 3.15	Rancangan Halaman Tambah <i>Stopwords</i>	39
Gambar 3.16	<i>Site Map</i>	40
Gambar 4.1	Halaman Utama	42
Gambar 4.2	Halaman Sistem Temu Kembali Model Ruang Vektor ..	43
Gambar 4.3	Isi File Pada <i>Dialogbox</i>	44
Gambar 4.4	Halaman Sistem Temu Kembali OKAPI BM25	46
Gambar 4.5	Isi File Pada <i>Dialogbox</i>	47
Gambar 4.6	Halaman Login	49

Gambar 4.7	Halaman Tambah Dokumen	50
Gambar 4.8	Halaman <i>Tambah Stopwords</i>	53
Gambar 4.9	Grafik Presisi-Recall Hasil Pengujian Kedelapan	64
Gambar 4.10	Grafik <i>Interpolated Precision</i>	65
Gambar 4.11	Grafik Presisi-Recall Hasil Pengujian Kedelapan	76
Gambar 4.12	Grafik <i>Interpolated Precision</i>	77
Gambar 4.13	Grafik Hasil Rata-rata Presisi-Recall Sistem Temu Kembali Model Ruang Vektor dan Sistem Temu Kembali OKAPI BM25	79

© UKDWN

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Seiring dengan perkembangan teknologi informasi yang semakin hari semakin pesat, manusia semakin dituntut untuk mendapatkan informasi secara cepat, tepat, dan akurat. Informasi yang dibutuhkan terkadang berupa dokumen, dan terkadang dokumen tersebut ada dalam jumlah yang sangat banyak sehingga pencarian secara manual hampir tidak mungkin dilakukan karena kemampuan otak manusia terbatas dan mudah sekali lupa. Dan akhirnya sebuah sistem telah ditemukan oleh para ahli untuk menyelesaikan permasalahan tersebut, sistem tersebut biasa disebut sistem IR (*Information Retrieval*).

Seiring berjalannya waktu metode yang digunakan untuk membangun sistem IR semakin bervariasi. Salah satunya adalah metode OKAPI BM25, metode ini merupakan metode probabilistik retrieval yang sangat terkenal karena mempunyai pembobotan yang baik. Begitu juga dengan metode Model Ruang Vektor (*Vector Space Model*), metode termasuk metode *ranked* retrieval dan mempunyai efektifitas kerja yang optimal sehingga masih banyak orang yang menggunakan sampai sekarang.

Dengan melihat kelebihan dari kedua metode tersebut penulis ingin mencoba membandingkan implementasi dari metode OKAPI BM25 dan metode Model Ruang Vektor pada sistem temu kembali.

1.2 Rumusan Masalah

Yang membuat penulis merasa tertarik untuk melakukan penelitian ini yang pertama adalah bagaimana hasil perbandingan dari sistem temu kembali yang mengimplementasikan metode OKAPI BM25 dengan sistem temu kembali yang mengimplementasikan metode Model Ruang Vektor. Dari kedua sistem tersebut, sistem mana yang memiliki performa yang lebih baik adalah pertanyaan yang ingin dijawab oleh penulis dalam penelitian ini.

Adapun yang performa yang baik diukur berdasarkan tingkat presisi serta *recall* yang dihasilkan dan kecepatan pemrosesan dari sistem untuk mengembalikan hasil *query*.

1.3 Batasan Masalah

Pada penelitian ini, pembuatan program atau sistem dibatasi oleh parameter-parameter sebagai berikut :

1.3.1 Format input

Format *query* berupa string dan dioperasikan implisit menggunakan oprand OR. Input query dipisahkan berdasar spasi. Sistem juga tidak dapat melakukan operasi khusus pada query yang berada didalam tanda petik , misalkan **“reaktor nuklir” jepang**. Sistem akan memperlakukan query tersebut sama seperti query biasa tanpa melakukan operasi khusus pada kata “reaktor nuklir”

Koleksi Dokumen berupa .txt atau .doc, dan tidak dapat mengenali format dokumen dalam format .pdf, .ppt, .xls atau yang lainnya.

1.3.2 Format output

Format output berupa string yang menampilkan list dari ranking dokumen, yang dihasilkan. Peringkat 1 berarti dokumen tersebut memiliki kemiripan yang paling besar diantara yang lain. Informasi yang dikembalikan oleh sistem tidak termasuk sinonim dari kata kunci yang diberikan oleh pengguna. Misalkan pengguna memasukkan kata 'usia' maka sistem hanya akan menggunakan kata usia saja dan tidak mempertimbangkan kata 'umur', meskipun usia dan umur memiliki makna yang sama.

1.4 Hipotesis

Metode Model Ruang Vektor yang melihat hubungan *query* dengan dokumen sedangkan pembobotan OKAPI BM25 tidak memperhatikan hubungan tersebut, Berdasarkan sifat tersebut maka penulis memiliki hipotesis bahwa sistem temu kembali yang mengimplementasikan metode Model Ruang Vektor memiliki tingkat presisi yang lebih tinggi.

1.5 Tujuan

Tujuan dari penelitian yang penulis lakukan ini adalah mengetahui metode mana yang memiliki performa lebih baik, apakah metode Okapi BM25 atau metode Model Ruang Vektor.

1.6 Metode Penelitian

Metodologi yang digunakan dalam penelitian ini adalah:

1. Studi pustaka dan literatur.

Studi pustaka dilakukan dengan mencari dan mempelajari sumber-sumber pustaka yang berkaitan dengan metode OKAPI BM25 dan metode VSM (*Vector Space Model*). Sumber-sumber tersebut berupa buku-buku, serta sumber *online* di Internet yang dapat dipercaya.

2. Implementasi

Penulis akan membangun 2 sistem temu kembali, dimana salah satunya mengimplementasikan metode OKAPI BM25 dan salah satunya mengimplementasikan metode Model Ruang Vektor.

3. Evaluasi

Evaluasi akan dilakukan oleh penulis dengan menghitung tingkat presisi dan *recall* dari masing-masing sistem.

1.7 Sistematika Penulisan

Sistematika laporan kerja praktek ini secara garis besar dapat dituliskan sebagai berikut:

BAB 1 merupakan PENDAHULUAN. Bab ini memuat latar belakang masalah, batasan masalah, metode penelitian, hipotesis, tujuan penelitian, dan sistematika penulisan.

BAB 2 merupakan TINJAUAN PUSTAKA. Tinjauan pustaka terdiri dari dua bagian utama yaitu , tinjauan pustaka dan landasan teori. Tinjauan pustaka menguraikan berbagai teori yang didapatkan dari sumber pustaka. Landasan teori memuat konsep dan penjelasan utama yang digunakan untuk memecahkan masalah dari penelitian yang dilakukan penulis.

BAB 3 merupakan ANALISIS DAN PERANCANGAN SISTEM.Memuat penjelasan tentang bahan/materi yang digunakan untuk melakukan riset, variable yang digunakan dalam penelitian dan cara perancangan sistem yang dibuat.

BAB 4 merupakan IMPLEMENTASI DAN ANALISIS SISTEM.Memuat hasil riset/implementasi dan pembahasan/analisi dari penelitian yang dilakukan.

BAB 5 merupakan KESIMPULAN DAN SARAN.Memuat kesimpulan dan saran dari penulis setelah melakukan penelitian.



BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Setelah dilakukan analisis dan evaluasi terhadap kedua sistem temu kembali yang telah dibangun, penulis menarik beberapa kesimpulan sebagai berikut :

1. Dengan query yang telah ditentukan oleh penulis serta jumlah dokumen dalam korpus sebanyak 105 dokumen kedua sistem mampu mengembalikan seluruh dokumen relevan yang terdapat dalam korpus (*Recall* 100%).
2. Karena penulis tidak melakukan proses *query parsing* dan hanya memisahkan query berdasarkan spasi, kualitas dari kedua sistem temu kembali menjadi kurang baik.
3. Pengukuran kecepatan perankingan dari kedua sistem tidak dilakukan, karena jumlah koleksi dokumen yang terlalu sedikit. Sehingga perhitungan tingkat kecepatan akan tidak signifikan.
4. Menilik dari rata-rata presisi yang dihasilkan, sistem temu kembali OKAPI BM25 mempunyai performa yang lebih baik dari pada sistem temu kembali Model Ruang Vektor.

5.2 Saran

Untuk meningkatkan hasil pencarian dan presisi dari kedua sistem dapat dilakukan dengan beberapa cara :

1. Menggunakan daftar sinonim pada proses indexing dan proses perangkingan. Dengan mendaftarkan kata umur sebagai sinonim kata usia atau Jogja dengan Jogjakarta misalnya, maka pada proses indexing token tersebut akan dianggap sama. Dan pada proses perangkingan sistem juga akan melakukan pencarian pada dokumen yang mengandung kata usia sekalipun pengguna memberikan kata kunci umur.

© UKDWN

DAFTAR PUSTAKA

- Grossman, A.D., & Frieder, O.(2004).*Information Retrieval:Algorithms and Heuristics*.Dordrecht:Springer.
- Hariyono,M.E.A.,&Wahyudi.(2005).*Customer Information Gathering Menggunakan Temu Kembali Informasi dengan Metode Ruang Vektor*.Diakses tanggal 15 Februari 2010 dari
<<http://journal.uui.ac.id/index.php/Snati/article/viewFile/1377/1157>>
- Kusrini.(2005).*Pengindekan dan Pencarian Dokumen Text*.Diakses tanggal 15 Februari 2010 dari
<<http://p3m.amikom.ac.id/p3m/dasi/sept06/05%20-%20STMIK%20AMIKOM%20Yogyakarta%20Makalah%20KUSRINI.pdf>>
- Manning, C.P., Raghavan, P., & Schütze, H.(2008).*Introduction to Information Retrieval*.USA:Canbrige University Press.
- Robertson, S.E., Walker, S., & Beaulieu, M.(2000).*Experimentation as a Way of Life: OKAPI at TREC*.Diakses tanggal 10 Februari 2011 dari
<<http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/ipmRobertson2000.pdf>>