

PENERAPAN SISTEM TEMU KEMBALI INFORMASI PADA KUMPULAN DOKUMEN SKRIPSI

Karter D. Putung, Arie Lumenta, Agustinus Jacobus
Teknik Informatika Universitas Sam Ratulangi Manado, Indonesia.
karterputung@gmail.com, al@unsrat.ac.id, a.jacobus@unsrat.ac.id

Abstrak - Sistem temu kembali informasi (*information retrieval system*) merupakan sistem yang digunakan untuk menemukan informasi yang relevan dengan kebutuhan dari pengguna, dengan menerapkan sistem tersebut permasalahan pencarian informasi dokumen skripsi bisa memberikan hasil yang relevan sesuai kebutuhan pengguna. Terdapat dua proses utama dalam sistem temu kembali informasi yaitu *indexing* dan *retrieval*. Proses *indexing* adalah proses untuk memberikan bobot pada kata dalam dokumen, metode pembobotan pada penelitian ini menggunakan metode pembobotan TF-IDF. Proses *retrieval* adalah proses untuk menghitung kemiripan *query* terhadap dokumen, perhitungan kemiripan menggunakan konsep *vector space model* dengan mencari nilai *cosine similarity*. Tujuan dari penelitian ini adalah untuk mengembangkan dan mengimplementasikan pengindeksan otomatis untuk membangun sistem pencarian dokumen di dalam sebuah system penyimpanan dokumen dengan konsep temu-kembali informasi.

Kata kunci : *Information retrieval, Term Frequency Inverse Document Frequency, Vector Space Model.*

I. PENDAHULUAN

Dokumen adalah contoh informasi yang tidak terstruktur, isi suatu dokumen sangat tergantung pada pembuat dokumen tersebut. Universitas Sam Ratulangi memiliki banyak data dokumen skripsi dari lulusan-lulusan mahasiswa Universitas Sam Ratulangi, dokumen-dokumen skripsi tersebut dapat digunakan oleh siapa saja untuk keperluan masing-masing.

Namun untuk melakukan pencarian dokumen skripsi sesuai dengan kebutuhan pencari informasi sangat sulit dilakukan, karena banyaknya dokumen yang ada. Untuk itu dibutuhkan sistem pencari informasi di dalam dokumen skripsi untuk memudahkan pencarian informasi yang dibutuhkan, dalam penelitian ini data dokumen skripsi yang akan digunakan dalam pembuatan sistem pencarian informasi hanya mengambil isi bagian abstrak dari dokumen skripsi yang akan di indeks.

Perbedaan antara temu kembali informasi dan pencarian data adalah, sistem temu kembali informasi terutama berhubungan dengan pencarian informasi yang isinya tidak memiliki struktur, yang kemudian ditemukan hasil suatu informasi yang dibutuhkan atau relevan dengan kata kunci.

Sifat pencarian sistem temu-kembali informasi berbeda dengan sistem temu-kembali data (misalnya dalam sistem manajemen basis data) dalam beberapa segi, antara lain

spesifikasi query yang tidak lengkap, dan tingkat ketanggapan kesalahan yang tidak peka [1].

Sistem temu-kembali data tujuan utamanya untuk menentukan dokumen yang memiliki kata kunci yang sesuai dengan query yang diberikan oleh pengguna di dalam sekumpulan dokumen, sehingga saat melakukan pencarian data jika data tidak sama persis maka informasi yang relevan dengan kata kunci tidak berhasil ditemukan. Hal ini belum dapat memecahkan masalah pengguna akan kebutuhan informasi.

II. LANDASAN TEORI

A. Information Retrieval

Information retrieval adalah sebuah proses untuk menemukan kembali informasi yang dibutuhkan dari sebuah sistem penyimpanan dan penelusuran informasi. Sistem temu Kembali Informasi mensyaratkan ada kebutuhan informasi dari pengguna, ada dokumen atau data yang berisi informasi yang diorganisasikan dalam sebuah sistem yang memudahkan temu kembali informasi dan strategi penelusuran yang tepat sehingga dokumen yang sesuai dengan kebutuhan dapat ditemukankembali [2].

Sistem temu balik informasi terdiri dari tiga komponen utama, yaitu masukan (input), pemroses (processor), dan keluaran (output). Menurut Manning et al (2009) Information Retrieval (IR) adalah menemukan bahan (biasanya dokumen) yang bersifat tidak terstruktur (biasanya teks) yang memenuhi kebutuhan informasi dari dalam koleksi besar (biasanya disimpan di komputer) [3].

B. Indexing

Indeks adalah bahasa yang digunakan di dalam sebuah buku konvensional untuk mencari informasi berdasarkan kata atau istilah yang mengacu ke dalam suatu halaman. Dengan menggunakan indeks si pencari informasi dapat dengan mudah menemukan informasi yang diinginkannya. Pada sistem temu-kembali informasi, indeks ini nantinya yang digunakan untuk merepresentasikan informasi di dalam sebuah dokumen [5].

Adapun tahapan dari pengindeksan adalah sebagai berikut:

1. *Pharsing* Memindahkan orang, baik di dalam maupun di luar batas Negara (termasuk perekrutan, pengangkutan, penampungan, pengiriman, pemindahan atau penerimaan).
2. *Stoptlist* yaitu proses pembuangan kata buang seperti: tetapi, yaitu, sedangkan, dan sebagainya.
3. *Stemming* yaitu proses penghilangan/ pemotongan dari suatu kata menjadi bentuk dasar.
4. *Term weighting* dan *Inverted File* yaitu proses pemberian bobot pada istilah.

C. Pembobotan TF-IDF

Metode TF-IDF (*Term Frequency Inverse Document Frequency*) merupakan suatu cara untuk memberikan bobot hubungan suatu kata (*term*) terhadap dokumen. Metode ini menggabungkan dua konsep untuk perhitungan bobot yaitu, frekuensi kemunculan sebuah kata didalam sebuah dokumen tertentu dan inverse frekuensi dokumen yang mengandung kata tersebut [4].

Formula yang digunakan Pada *Term Frequency (tf)*, terdapat yaitu nilai *tf* diberikan berdasarkan jumlah kemunculan suatu kata di dokumen. Contohnya, jika muncul lima kali maka kata tersebut akan bernilai lima.

Inverse Document Frequency (idf) dihitung dengan menggunakan formula sebagai berikut [3] :

$$IDF = \log\left(\frac{D}{DF}\right) \quad (1)$$

Dimana,

D = Jumlah dokumen.

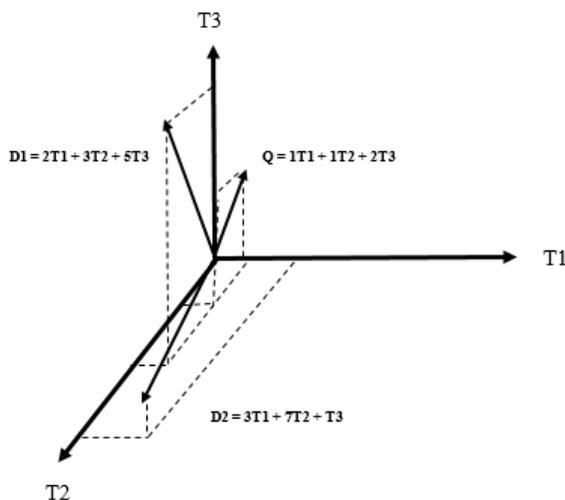
DF = Dokumen frekuensi.

Dengan demikian rumus umum untuk TF-IDF adalah penggabungan dari formula perhitungan *raw tf* dengan formula *idf* dengan cara mengalikan nilai *term frequency (tf)* dengan nilai *inverse document frequency (idf)* [3]:

$$W = tf \times IDF \quad (2)$$

D. Vector Space Model

Metode *Vector Space* adalah sebuah model aljabar untuk menggambarkan dokumen teks (beberapa objek) sebagai vektor dari identifier. Biasanya digunakan dalam penyaringan informasi (*information filtering*), penemuan informasi (*information retrieval*), *indexing* dan pemberian ranking yang saling relevan [5].



Gambar 1. Reperentasi dokumen dan query pada ruang vektor

Model ruang vektor mengasumsikan bahwa sudah ada index terms yang telah merepresentasikan dokumen dan query. Sebuah dokumen D_i dan Q_j query direpresentasikan seperti berikut ini [6] :

$$D_i = [T_{i1}, T_{i2}, \dots, T_{ik}, \dots, T_{iN}]$$

$$Q_j = [Q_{j1}, Q_{j2}, \dots, Q_{jk}, \dots, Q_{jN}]$$

dimana T_{ik} adalah nilai dari *termk* dalam dokumen i , Q_{jk} adalah nilai *queryk* dalam *query* j , dan N adalah jumlah total istilah yang digunakan dalam dokumen dan *query*. Nilai dari T_{ik} dan Q_{jk} dapat di ambil dari hasil pembobotan perhitungan TF-IDF [6].

Pencarian dalam model ruang vektor didasarkan pada kesamaan antara query dan dokumen. Kesamaan antara dokumen D_i dan permintaan Q_j koefisien kesamaan antara query dan dokumen dapat diperoleh dengan formula inner product.

$$S(D_i, Q_j) = \sum_{k=1}^n (T_{ik} \cdot Q_{jk}) \quad (3)$$

Namun dengan formula 3 dokumen yang lebih panjang dengan jumlah istilah yang lebih banyak memiliki kemungkinan lebih besar untuk dianggap relevan dengan query tertentu dibandingkan dokumen-dokumen yang lebih pendek. Sehingga formula di atas digantikan dengan formula ternormalisasi [5] :

$$S(D_i, Q_j) = \frac{\sum_{k=1}^N T_{ik} \cdot Q_{jk}}{\sqrt{\sum_{k=1}^N T_{ik}^2 \cdot \sum_{k=1}^N Q_{jk}^2}} \quad (4)$$

Nil

ai cosinus yang cenderung besar mengindikasikan bahwa dokumen cenderung sesuai query. Nilai cosinus sama dengan 1 mengindikasikan bahwa dokumen sesuai dengan query.

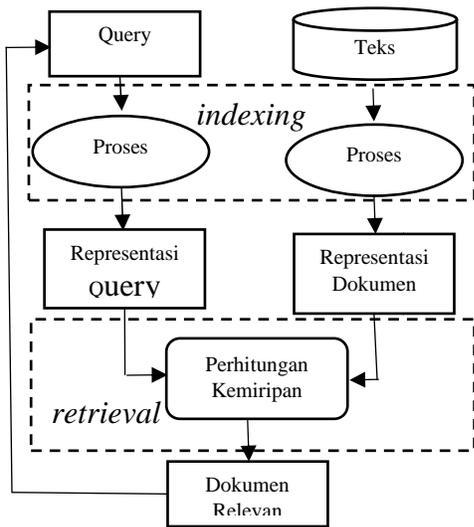
III. METODOLOGI PENELITIAN

A. Analisis Sistem

Analisis sistem dilakukan untuk menganalisa sistem temu kembali informasi yang akan dibangun sebelum diimplementasikan adapun analisa yang akan dilakukan yaitu analisis proses sistem temu kembali informasi secara keseluruhan, kemudian menganalisa proses indexing dan proses retrieval.

Proses utama dari sistem temu kembali informasi yang dibangun terbagi menjadi 2 bagian yang pertama adalah query, tahap pertama query akan diproses menjadi term untuk direpresentasikan ke tahap selanjutnya, representasi dari query adalah token atau term yang sudah memiliki nilai. Bagian kedua adalah dokumen, tahapan dari dokumen sama dengan tahapan pada query. Setelah query dan dokumen sudah

direpresentasikan maka akan dilakukan tahap perhitungan kemiripan, sehingga dokumen yang relevan dengan query ditemukembali

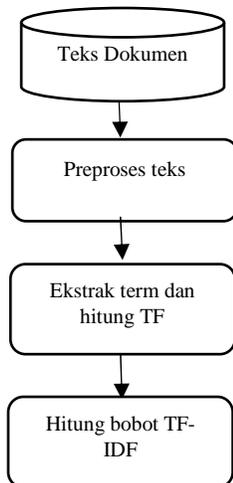


Gambar 2. Proses sistem temu kembali informasi

1. Proses Index

Proses index adalah proses untuk memberikan nilai bobot dari setiap term pada dokumen dan menghitung nilai bobot dari dokumen kemudian di simpan ke dalam basis data tindex. Proses index ini terbagi dari beberapa proses yaitu proses preproses teks, proses ekstrak term dan hitung tf, proses hitung bobot TF-IDF.

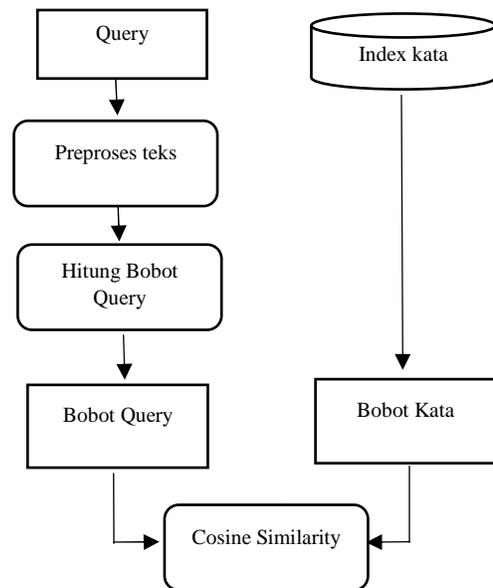
- Proses preproses teks adalah proses untuk menghilangkan tanda baca dan merubah semua huruf menjadi huruf kecil pada data abstrak skripsi.
- Proses ekstrak term dan hitung tf, adalah proses untuk mengambil setiap kata dari abstrak skripsi kemudian menghitung frekuensi kemunculan kata dan meyimpannya kedalam basis data.
- Proses hitung bobot TF-IDF adalah proses untuk menghitung nilai bobot dari term yang ada pada basis data.



Gambar 3. Proses indeks

2. Proses retrieval

Proses hitung kemiripan adalah proses untuk mencari kemiripan antara dokumen dengan kata kunci dengan mencari nilai cosine similarity.



Gambar 4. Proses retrieval

B. Perancangan Perangkat Lunak

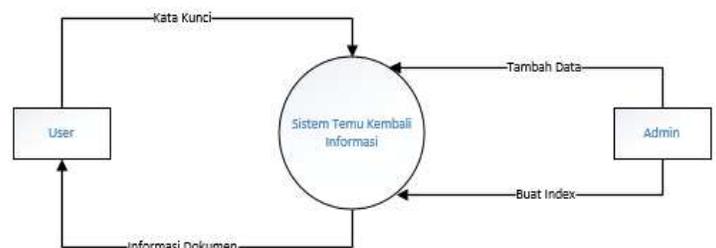
Tahap perancangan adalah proses penggunaan berbagai teknik dan prinsip untuk tujuan mendefinisikan proses atau sistem secara detail. Adapun tahapan dari perancangan perangkat lunak adalah :

1. Rencana kebutuhan

Rencana kebutuhan dalam penelitian ini secara garis besar sistem temu kembali informasi yang akan dibangun memiliki dua proses utama yaitu proses *indexing* dan proses *query (retrieval)*. Proses *indexing* yang berhubungan dengan dokumen-dokumen sedangkan proses *query (retrieval)* yang berhubungan dengan pengguna.

2. Pemodelan proses

Pemodelan proses pada penelitian ini dilakukan dengan menggunakan tools Data Flow Diagram (DFD), Data Flow Diagram adalah alat yang menggambarkan aliran data melalui sistem dan kerja/tugas atau pengolahan yang dilakukan oleh system tersebut



Gambar 5. DFD Level 0

DFD Level 0, adalah sebuah teknik grafik yang menggambarkan aliran data dari keseluruhan sistem. Dengan melihat gambar 5 DFD Level 0 menjelaskan bahwa sistem yang akan dibangun memiliki dua aktor yaitu user dan admin.

3. Perancangan basis data

Perancangan Basis Data adalah proses untuk menentukan isi dan pengaturan data yang dibutuhkan untuk mendukung berbagai rancangan sistem. Adapun Basis data dan struktur table yang akan dibuat pada sistem temu kembali informasi adalah table tbskripsi, tbindeks, tbvektor, dan tbretrieval.

4. Perancangan antarmuka

Pada tahap ini antarmuka dari perangkat lunak akan dibuat dalam bentuk storyboard untuk visualisasi ide dari aplikasi yang akan dibangun, sehingga dapat memberikan gambaran dari aplikasi yang akan dihasilkan.

IV. HASIL DAN PEMBAHASAN

A. Implementasi sistem

Tahap implementasi menerjemahkan hasil analisa dan perancangan sehingga menghasilkan suatu sistem temu kembali informasi yang dapat digunakan pengguna untuk mencari informasi dokumen skripsi.

1. Proses indeks.

Implementasi dari proses indeks adalah proses untuk memberikan nilai bobot dari setiap term pada dokumen dan kemudian di simpan ke dalam basis.

Dokumen skripsi yang di indeks yang menjadi sampel data dalam penelitian ini sebanyak 100 dokumen. Beberapa data hasil dari proses tersebut dapat dilihat pada tabel berikut :

Tabel 1. Hasil perhitungan indeks.

No	Term	Dok_Id	TF	TF-IDF
1	Sistem	1	6	1,42649
		2	1	0,237748
		3	2	0,475496
		6	2	0,475496
		7	2	0,475496
2	Informasi	4	6	3,18887
		9	2	1,06296
		11	9	4,78331
		14	3	1,59444
3	Aplikasi	15	3	1,59444
		9	2	1,22132
		11	1	0,61066
		14	3	1,83198
		15	2	1,22132
4	Android	24	4	2,44264
		14	3	3,0258
		24	1	1,0086
		28	1	1,0086
5	Animasi	47	5	5,043
		5	7	9,84578
		31	7	9,84578
		38	3	4,21962
6	Camera	59	4	5,62616
		9	2	2,81308
		16	1	1,40654
		86	4	5,62616

		93	2	2,8130
7	Perancangan	3	1	0,862472
		4	1	0,862472
		29	2	1,72494
		37	1	0,862472

Dimana,

Term = Kata TF = Frekuensi kata

Dok_Id = Id dokumen TF-IDF = Nilai Bobot

Data yang terdapat dalam tabel 1 hasil perhitungan indeks adalah sebagian data yang diperoleh dari hasil proses indeks yang dilakukan sistem, misalnya untuk kata "sistem" yang ada pada dok_id 1 berjumlah 6 dengan bobot perhitungan berdasarkan formula 2 perhitungan TF-IDF bernilai 1,42649.

2. Proses retrieval

Implementasi dari proses *retrieval* adalah proses untuk mencari dokumen skripsi yang relevan berdasarkan kata kunci yang di masukan pengguna.

Hasil dari implementasi proses retrieval dapat dilihat pada tabel berikut.

Tabel 2. Hasil pencarian sistem.

No	Kata Kunci	Dok_Id	Kemiripan
1	Augmented reality	17	0,142223
		59	0,425442
2	Tourism Kuliner	24	0.354562
		45	0.116486
		58	0.171479
3	Render animasi	5	0.307896
		31	0.307896
		38	0.254763
		59	0.134094
4	Penunjuk arah jalan	16	0.0964674
		19	0.10385
		22	0.0671116
		51	0.0328305
		83	0.21451
		91	0.128767
5	Ip camera	9	0.134485
		16	0.0489398
		22	0.156334
		46	0.040032
		86	0.201221
		92	0.258416
		93	0.174418

Hasil pencarian dari sistem yang ada dalam tabel 2 hasil pencarian adalah beberapa data implementasi dari hasil pencarian dokumen berdasarkan kata kunci yang dimasukkan pengguna dengan mencari nilai cosine similarity berdasarkan formula 4. Dapat dilihat untuk kata kunci "augmented reality", sistem menemukan 2 dokumen yang relevan yaitu Dok_Id 17 dengan kemiripan 0,142223 dan Dok_Id 59 dengan kemiripan 0,425442.

B. Pengujian Indexing

Pengujian ini dilakukan untuk mengetahui waktu pembuatan index pada sistem, waktu pembuatan index akan diuji berdasarkan banyaknya dokumen yang di index dan lamanya proses index. Sample data yang dimasukan sebanyak 100 dokumen dan akan dihitung waktu proses index setiap 10 dokumen.

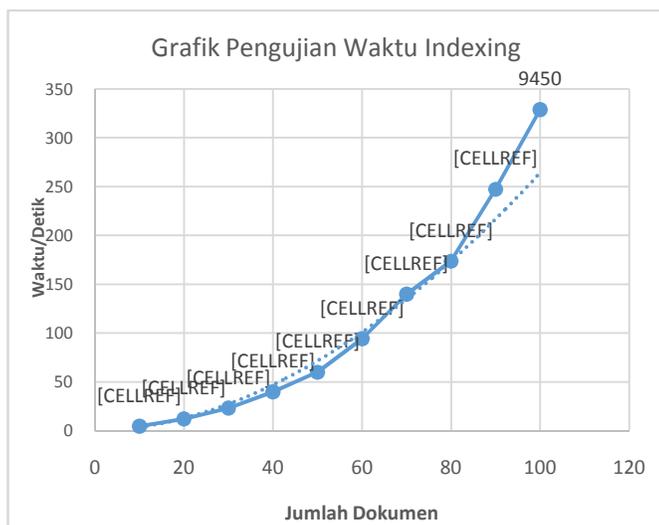
Pengujian *indexing* sebanyak 100 dokumen lama proses *indexing* yang dibutuhkan sistem yaitu 328.84 detik dengan jumlah term sebanyak 9480 *term* yang di berikan bobot.

Untuk melihat hasil keseluruhan pengujian waktu *indexing* yang telah dilakukan dapat dilihat pada Tabel 3 pengujian waktu *indexing*.

Tabel 3 pengujian waktu *indexing*.

No	Banyak Dokumen	Waktu index / Detik	Keterangan
1	10	4,41	1043 <i>Term</i>
2	20	11,89	1988 <i>Term</i>
3	30	22,96	2896 <i>Term</i>
4	40	39,72	3996 <i>Term</i>
5	50	59,76	5036 <i>Term</i>
6	60	94,09	6025 <i>Term</i>
7	70	139,72	7064 <i>Term</i>
8	80	173,48	7680 <i>Term</i>
9	90	247,18	8579 <i>Term</i>
10	100	328,85	9450 <i>Term</i>

Dapat dilihat pada gambar 6 grafik pengujian indexing, bahwa waktu proses *index* terus meningkat bersama dengan banyak dokumen dan jumlah *term* yang di *index*. Tidak menutup kemungkinan hasil pengujian akan berbeda jika menggunakan spesifikasi komputer yang lain, atau dengan versi perangkat lunak yang lain.



Gambar 6. grafik pengujian *indexing*

C. Pengujian Retrieval

Pengujian sistem untuk proses *retrieval* akan dilakukan dengan menguji dua parameter utama yaitu *recall* dan *precision*.

1. Pengujian *recall*

Recall adalah rasio antara dokumen yang relevan yang berhasil ditemukembalikan dari seluruh dokumen relevan yang ada didalam sistem nilai tertinggi *recall* adalah 1 yang berarti seluruh dokumen dalam koleksi berhasil ditemukembalikan, *recall* didefinisikan sebagai berikut :

$$recall = \frac{\text{jumlah dokumen relevan ditemukan}}{\text{jumlah dokumen relevan dalam koleksi}} \quad (5)$$

2. Pengujian *precision*

Precision adalah rasio dokumen relevan yang berhasil ditemukembalikan dari seluruh dokumen yang ditemukembalikan nilai tertinggi *precision* adalah 1 yang berarti seluruh dokumen yang ditemukan adalah relevan, *precision* didefinisikan sebagai berikut :

$$precision = \frac{\text{jumlah dokumen relevan ditemukan}}{\text{jumlah dokumen ditemukan}} \quad (6)$$

Hasil pengujian *recall* dan *precision* pada sistem dengan menggunakan data sampel sebanyak 100 dokumen. Diperlihatkan pada tabel 4 pengujian *recall* dan *precision*. Pengujian dilakukan sebanyak 10 kali, dengan menggunakan kata kunci yang berbeda di setiap pengujian.

Tabel 4 pengujian *recall* dan *precision*.

No	Pengujian	Kata kunci	Ra	Rs	Rt	Recall	Precision
1	T1	Augmented Reality	2	2	2	1	1
2	T2	tourism kuliner	3	3	3	1	1
3	T3	render animasi	3	3	4	1	0,75
4	T4	penunjuk arah jalan	4	4	6	1	0,666667
5	T5	ip camera	3	3	7	1	0,428571
6	T6	aplikasi android	9	9	26	1	0,346154
7	T7	Newton-Rapshon	2	2	2	1	1
8	T8	Penunjuk arah jalan ip camera	7	7	11	1	0,636364
9	T9	Render animasi tourism kuliner	6	6	7	1	0,857143
10	T10	Augmented Reality ip camera	5	5	9	1	0,5

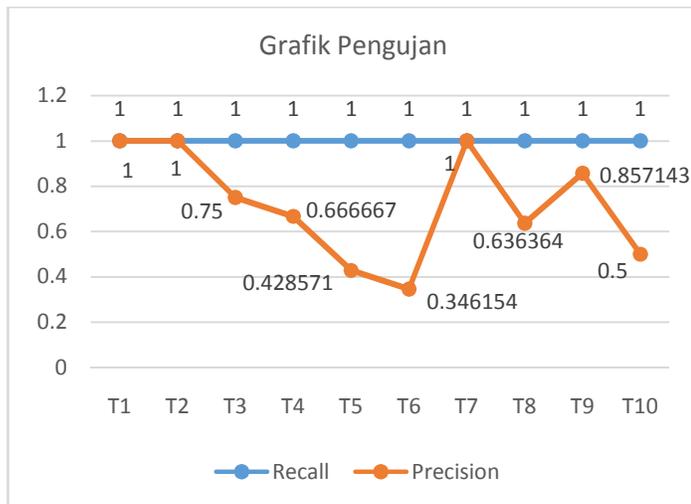
Dimana,

Ra = jumlah dokumen relevan ditemukan

Rs = jumlah dokumen relevan dalam koleksi

Rt = jumlah dokumen ditemukan

Untuk melihat grafik hasil pengujian *retrieval* dapat dilihat pada gambar 7 grafik *recall* dan *precision*.



Gambar 7. grafik pengujian *recall* dan *precision*

Grafik *recall* membentuk garis lurus dengan nilai 1 yang artinya seluruh dokumen dalam koleksi berhasil di temukembalikan. Nilai dari *precision* memiliki nilai yang berbeda hampir di setiap pengujian, dapat dilihat dari garis grafik yang naik turun.

Nilai dari *precision* tergantung dari keunikan kata kunci yang diberikan, semakin unik kata kunci yang diberikan maka semakin tinggi nilai *precision* yang diperoleh sebaliknya semakin umum kata kunci yang diberikan maka akan semakin kecil nilai dari *precision* tersebut.

V. PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan sistem temu kembali informasi berhasil dibangun dan dapat mengimplementasikan pengindeksan otomatis dengan menggunakan metode pembobotan TF-IDF dan dapat menemukan dokumen yang relevan terhadap kata kunci dengan menghitung nilai *cosine similarity*. Waktu proses *indexing* sistem berdasarkan pengujian waktu *indexing*, waktu proses *indexing* terus meningkat bersamaan dengan banyaknya dokumen yang di *index* dengan waktu pengujian *indexing* pada 100 dokumen skripsi 328.85 detik.

Data hasil pengujian menghasilkan nilai *recall* 1 pada setiap pengujian yang menunjukkan bahwa semua dokumen yang relevan dapat ditemukan oleh sistem, nilai *precision* antara 0.346154 sampai dengan 1 yang menunjukkan terdapat dokumen lain selain dokumen relevan yang ikut ditemukan oleh sistem.

B. Saran.

Beberapa saran yang diusulkan untuk penelitian dan pengembangan sistem temu kembali informasi yaitu :

1. Dikembangkan algoritma *stemming* (penghilangan imbuhan) untuk kata-kata berbahasa Indonesia. Penggunaan *stemming* dapat mengurangi ukuran dari *indeks* pada sistem temu-kembali informasi dan Penggunaan *thesaurus* (pengelompokan kata yang berhubungan) sehingga dapat meningkatkan *precision* dari hasil pencarian temu-kembali informasi.
2. Penggunaan model dari temu-kembali informasi yang lainnya untuk dapat membandingkan hasil kinerja temu-kembali informasi sehingga dapat ditemukan model yang paling baik dari sistem temu-kembali informasi.

DAFTAR PUSTAKA

- [1] Rijsbergen, C.J Van. 1979. Information Retrieval. Second Edition. Butterworths, London.
- [2] Kurniasi, N. 2014. Konsep Dasar *Information Retrieval System* / Sistem Temu Kembali Informasi. Materi Kuliah Departemen Ilmu Informasi Dan Perpustakaan Universitas Padjajaran.
- [3] Manning, D. Christopher, Raghavan, P. & Schütze H. 2009. *An Introduction to Information Retrieval*. Cambridge University Press
- [4] Karmayasa, O. 2012. Implementasi *Vector Space Model* / Metode *Term Frequency Inverse Document Frequency* (TF-IDF) Pada Sistem Temu Kembali Informasi. Universitas Udayana. Denpasar
- [5] Nadirman, F. 2006. Sistem Temu-kembali Informasi Dengan Metode *Vector Space Model* Pada Pencarian *File* Dokumen Berbasis Teks. Skripsi Program Studi Ilmu Komputer Universitas Gadjah Mada. Yogyakarta.
- [6] Lu, Goujun. 1963. *Multimedia Database Management Systems*. Artech House. Norwood



Sekilas tentang penulis dengan nama lengkap Karter Dagels Putung, lahir di desa Paslaten, Kecamatan tatapaan, Kabupaten Minahasa Selatan, Provinsi Sulawesi Utara. Anak ke-1 dari 4 bersaudara. Dengan pendidikan Sekolah Dasar, SDGIMWawona. Kemudian Melanjutkan ke Sekolah Menengah Pertama, SMP Negeri 1 Tumpaan. Setelah lulus dari Sekolah Menengah Pertama melanjutkan ke Sekolah Menengah Atas, SMAN 1 Amurang. Setelah lulus tahun 2007 melanjutkan pendidikan Perguruan Tinggi di Universitas Sam Ratulangi Manado pada tahun 2011 dengan mengambil Jurusan Teknik Informatika. Pada tahun 2015 bulan Agustus, penulis membuat Skripsi demi memenuhi syarat Sarjana (S1) dengan penelitian berjudul Penerapan Sistem Temu Kembali Informasi Pada Kumpulan Dokumen Skripsi yang dibimbing oleh dua dosen pembimbing yaitu Arie S.M Lumenta ST., MT dan Agustinus Jacobus ST., M.Cs sehingga pada tanggal 6 April 2016 penulis resmi lulus di Teknik Informatika Universitas Sam Ratulangi Manado dan menyandang gelar Sarjana Komputer.