



## Penerapan Model Gravitasi Newton Versi Continuous dan Diskrit pada Sistem Temu Balik Informasi

Benie Ilman<sup>a</sup>, Iwan Abadi<sup>b</sup>, Nezar Rizky Aulia<sup>c</sup>

<sup>a</sup>Teknik Informatika, Fakultas Teknik, Unla, benie.ilman@gmail.com

<sup>b</sup>Teknik Informatika, Fakultas Teknik, Unla, iwan.abadi69@gmail.com

<sup>c</sup>Teknik Informatika, Fakultas Teknik, Unla,

### Abstract

The information retrieval system (IRS) is related to the retrieval of information from documents based on the context and content of the documents themselves. The process in IRS can be described as a process to get relevant documents from a collection document through search queries that users input. Test parameters to assess the relevance of a document used are precision and recall. Precision is a parameter to calculate the degree of certainty between the query and the collection document. Precision is the result of comparisons between documents relevant to all documents successfully retrieved by the system. Recall is a parameter to calculate the value of completeness between queries and collection documents. Recall is the result of comparison between relevant documents and relevant documents in the entire collection document. To get similarity value of document then in the process use Newton gravity model. Newton's gravity model is a representation of the magnetic field, the inverse square magnetic field ( $1/x^2$ ) and the negative exponential magnetic field ( $e^{-x}$ ), representing the continuous version, and the discrete weighting function as a representation of the discrete version.

*keywords:* information retrieval, precision, recall, gravitasi newton, continuous, discrete

### Abstrak

Sistem temu balik informasi (STBI) adalah yang berhubungan dengan pengambilan informasi dari dokumen-dokumen yang didasarkan pada isi dan konteks dari dokumen-dokumen itu sendiri. Proses dalam SBTI dapat digambarkan sebagai sebuah proses untuk mendapatkan dokumen yang relevan dari sekumpulan dokumen melalui pencarian *query* yang diinputkan user. Parameter uji untuk menilai relevansi sebuah dokumen yang digunakan adalah *precision* dan *recall*. *Precision* adalah parameter untuk menghitung nilai tingkat kepastian antara query dengan koleksi dokumen. *Precision* merupakan hasil dari perbandingan antara dokumen relevan dengan seluruh dokumen yang berhasil diambil oleh sistem. *Recall* adalah parameter untuk menghitung nilai tingkat kelengkapan antara query dengan dokumen koleksi. *Recall* merupakan hasil perbandingan antara dokumen relevan dengan dokumen relevan yang ada dalam seluruh dokumen koleksi. Untuk mendapatkan nilai *similarity* dokumen, maka dalam prosesnya menggunakan model gravitasi Newton. Model gravitasi Newton merupakan representasi dari medan magnet yaitu medan magnet inverse kuadrat ( $1/x^2$ ) dan medan magnet eksponensial negatif ( $e^{-x}$ ), yang mewakili versi continuous, dan fungsi pembobotan diskrit sebagai representasi dari versi diskrit.

*Kata Kunci:* temu balik informasi, precision, recall, gravitasi newton, continuous, diskrit

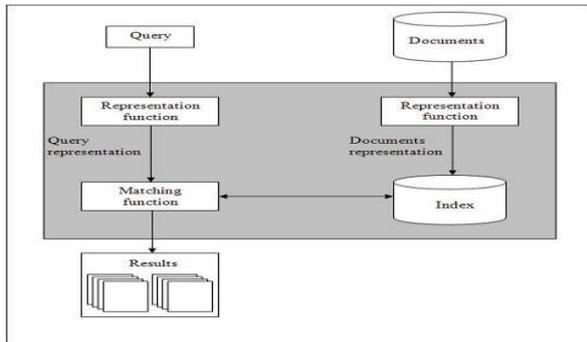
© 2017 Prosiding SISFOTEK

### 1. Pendahuluan

Sistem Temu Balik Informasi (SBI) adalah sistem pencarian informasi-informasi relevan dari sekumpulan dokumen yang didasari kebutuhan informasi pengguna yang di-input-kan ke sistem dalam bentuk *query*. *Query* merupakan kata atau kumpulan kata yang dijadikan acuan oleh sistem untuk mengembalikan dokumen yang sesuai kepada pengguna. STBI berkaitan dengan representasi, penyimpanan, dan akses terhadap dokumen representasi dokumen [1]. Parameter dalam

mengukur kinerja STBI secara umum menggunakan nilai Perolehan (*recall*) dan Ketepatan (*precision*). *Recall* merupakan nilai pecahan atau bagian dari dokumen relevan yang dikembalikan sistem berdasarkan total dokumen relevan yang ada di koleksi dokumen. *Precision* merupakan nilai pecahan dari dokumen yang relevan yang dikembalikan sistem berdasarkan total dokumen yang dikembalikan sistem. Tujuan umum dari STBI adalah menemu-balikkan informasi yang relevan dengan kebutuhan pengguna. Sistem informasi yang ideal adalah yang dapat

menemukan semua dokumen yang relevan dan tidak mengembalikan satu pun dokumen yang tidak relevan[2]. Gambar 1 di bawah menggambarkan prinsip sistem temu balik informasi.



Gambar 1 Prinsip Sistem Balik Temu Informasi

Dalam sistem temu balik informasi diperlukan pemodelan yang mampu menerjemahkan dokumen, query dan relasi yang relevan. Model yang umum digunakan pada Sistem Temu Balik Informasi diantaranya adalah *Boolean Models*, *Vector Space Models*, *Probabilistic and Logic Models*, dan *Language Models*. *Boolean Models* menggunakan pendekatan teori himpunan dan aljabar boolean, sedangkan *Vector Space Model* menggunakan pendekatan secara vektor dan aljabar linier. Pada model-model tersebut belum ada yang menerjemahkan sistem temu balik informasi melalui pendekatan representasi fisik. Sehingga pada beberapa dekade terakhir, model Gravitasi (*Model Gravitasi Newton*).

Sistem temu balik informasi melakukan 4 tahap utama sebelum dapat mengembalikan dokumen kepada pengguna meliputi [3]:

1. *Text Operations* (operasi terhadap teks) yang meliputi pemilihan kata-kata dalam *query* maupun dokumen (*term selection*) dalam pentransformasian dokumen atau *query* menjadi *terms index* (indeks dari kata-kata).
2. *Indexing* (pengindeksan), membangun basis data indeks yang lazim disebut *inverted file* dari koleksi dokumen.
3. *Query formulation* (formulasi terhadap *query*) yaitu memberi bobot pada indeks kata-kata *query*.

*Ranking* (perankingan), adalah mencari dokumen-dokumen yang relevan terhadap *query* dan mengurutkan dokumen tersebut berdasarkan kesesuaiannya dengan *query*.

## 2. Tinjauan Pustaka/Penelitian Sebelumnya

### 2.1 Pemodelan Sistem Temu Balik Informasi

Model Sistem Temu Balik Informasi adalah metode yang digunakan untuk melakukan pencocokan antara *terms* dari *query* dengan *terms* dalam koleksi dokumen. Ada 3 model besar, yaitu[4]

- a. *Set-theoretic models*, model merepresentasikan dokumen sebagai himpunan kata atau frase. Contoh model ini ialah *standard Boolean model* dan *extended Boolean model*.
- b. *Algebraic model*, model merepresentasikan dokumen dan *query* sebagai vektor atau matriks *similarity* antara vektor dokumen dan vektor *query* yang direpresentasikan sebagai sebuah nilai skalar. Contoh model ini ialah *vector space model* dan *latent semantic indexing (LSI)*.
- c. *Probabilistic model*, model memperlakukan proses pengembalian dokumen sebagai sebuah *probabilistic inference*. Contoh model ini ialah penerapan teorema *Bayes* dalam model probabilistik.

### 2.2 Model Gravitasi

Gravitasi mengatur gerakan bintang dan memegang peranan penting dari mayoritas proses di bumi. Model Gravitasi ini memetakan konsep hukum fisika Newton seperti massa, jarak, radius dan gaya tarik menarik[5]. Gambaran proses tersebut menjadi konsep sistem temu balik informasi seperti dokumen, *query* dan hubungan antara keduanya. Untuk memahami secara pasti esensi dan sudut pandang yang paling bisa menerjemahkan sistem temu balik informasi secara tepat dan menyeluruh sangatlah sulit. Namun pemahaman mengenai sistem temu balik informasi akan semakin baik apabila sebuah sudut pandang baru berhasil digunakan dan diadaptasi. Pemetaan STBI ke dalam model berbasis gravitasi ditunjukkan pada Tabel 1.

Tabel 1 Model berbasis Gravitasi Newton

Expression	Description
$m(t, D)$	Massa dari <i>term</i> t dalam dokumen D
$m(t, Q)$	Massa dari <i>term</i> t dalam <i>query</i> Q
$\bar{m}(t)$	Massa rata-rata <i>term</i> t pada koleksi dokumen
$m(D)$	Massa dari dokumen D
$di(t, D)$ $di(D)$	Diameter dari <i>term</i> t dalam dokumen D Diameter dari dokumen D
$r(t, D)$	Radius dari <i>term</i> t dalam dokumen D
$H(D)$	<i>Hidden terms</i> pada dokumen D

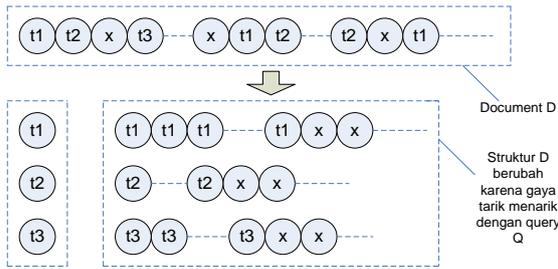
Dalam gravitasi newton, sebuah *term* yang terdapat dalam suatu dokumen ditinjau sebagai suatu objek fisik yang terdiri dari sejumlah partikel. Partikel memiliki dua atribut yaitu jenis dan massa. Jenis suatu partikel ditentukan oleh *term* yang menyusunnya. Dua partikel dengan jenis yang sama memiliki suatu gaya gravitasi antar keduanya sehingga kedua partikel tersebut akan saling tarik-menarik. Jika kedua partikel tersebut berbeda jenisnya, maka tidak akan terdapat suatu gaya antar keduanya.  $P_1, P_2$  adalah partikeldengan massa

sebesar  $m_1$  and  $m_2$ , maka gaya gravitasi antar keduanya adalah

$$F(P_1, P_2) = \frac{G.m_1.m_2}{d^2}, \text{ jika type } (P_1) = \text{type } (P_2) \quad (1)$$

### 2.3 Pemodelan Gravitasi Newton Versi Diskrit

Default dokumen merupakan kumpulan *term* yang disortir berdasarkan urutannya, tetapi struktur dokumen tersebut diasumsikan berubah-ubah sesuai dengan daya tarik/*attraction* dari *query* yang diberikan. Gambar 2 di bawah merupakan model gravitasi Newton versi diskrit'



Gambar 2 Model Gravitasi Newton Versi Diskrit

Dari gambar 2, lingkaran adalah *query terms* yang ada dalam dokumen yang diberi label “t1”, “t2”. Sedangkan bola yang diberi label “x” adalah *term* lain yang bukan bagian dari *query* Q. *Term* yang memiliki gaya gravitasi yang lebih besar akan semakin dekat dengan *query* yang diberikan. *Query* yang diberikan pada setiap dokumen memiliki *optimized termplacement*, dimana penjumlahan gaya yang ada antara *query* dengan dokumen yang bersangkutan adalah maksimal. Dalam model gravitasi, relevansi dari dokumen yang mengandung *query* didefinisikan oleh gaya tarik-menarik antar keduanya, dimana dokumen berada pada *optimal-term-placement state*. persamaan di bawah ini, bahwa gaya gravitasi yang optimal antara *query term* dan dokumen D adalah:

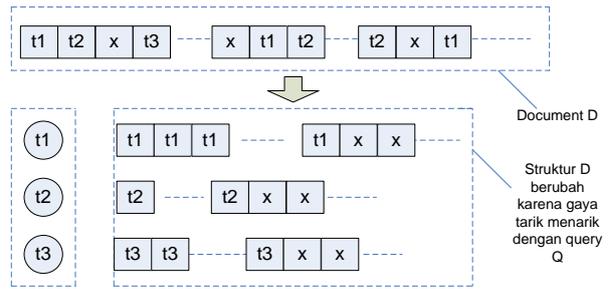
$$F_{gbm}(t, D) = \sum_{i=0}^{c(t,D)-1} \frac{G \times m(t, D)}{(l + (i + \frac{1}{2}) \times di(t, D))^2} \quad (2)$$

Berdasarkan prinsip komposisi gaya (*principle of the composition of forces*), gaya yang terjadi antara dokumen D dan *query* Q dinyatakan sebagai berikut

$$F_{gbm}(Q, D) = \sum_{t \in Q} F_{gbm}(t, D) \quad (3)$$

### 2.4. Pemodelan Gravitasi Newton Versi Continuous

Gambar 3 menggambarkan Pemodelan Gravitasi Newton Versi Continuous.



Gambar 3 Model Gravitasi Newton Versi Continuous

dapat diketahui bahwa gaya tarik menarik antara sebuah istilah *query* t dan dokumen D dapat dinyatakan sebagai berikut

$$F_{gbm}(t, D) = \int_l^{l+c(t,D).d(t,D)} \frac{G.m(t, D)}{di(t, D)x^2} dx = \frac{G.m(t, D)}{l.di(t, D)} \cdot \frac{c(t, D)}{\frac{l}{di(t, D)} + c(t, D)} \quad (4)$$

### 2.4 Perkiraan Perhitungan Massa dan Diameter

Sebuah dokumen memiliki massa dan panjang yang berbeda-beda. Sebuah *term* yang sama juga mungkin memiliki nilai massa yang berbeda-beda pada dokumen yang berbeda. Dua *term* yang sama: “lulus”, nilai massanya mungkin akan berbeda jika “lulus” tersebut berada pada dokumen “satu” dengan “lulus” yang ada pada dokumen “dua”. Meskipun massa dari sebuah *term* independen untuk setiapdokumen. Massa rata-rata dari sebuah *term* dalam seluruh himpunan dapat dinyatakan sebagai  $\bar{m}(t)$  [6].

Untuk menyederhanakan dan membuat proses menjadi lebih ringkas digunakan beberapa asumsi penyederhanaan yang menjadi dasar dalam melakukan perhitungan yang ada.

#### Asumsi 1:

Untuk setiap dua *term*, rasio perbandingan massa dalam dokumen manapun terhadap massa rata-rata di keseluruhan koleksi dokumen adalah sama, dinyatakan sebagai berikut:

$$\frac{m(t_1, D)}{m(t_1)} = \frac{m(t_2, D)}{m(t_2)} \xrightarrow{\text{(dinyatakan sebagai)}} h(D) \quad (5)$$

#### Asumsi 2:

Massa rata-rata dari suatu *term*, bergantung dan hanya bergantung pada frekuensi invers dokumen di dapatkan:

$$\bar{m}(t) = \ln\left(\frac{N+1}{df(t)}\right) \quad (7)$$

dimana N adalah total jumlah dari dokumen pada koleksi,  $df(t)$  adalah dokumen frekuensi dari *term* t.

**Asumsi 3:**

Rata-rata kepentingan global dari semua *term* pada suatu dokumen adalah konstan (independen dari dokumen itu sendiri). Asumsi ini mengimplikasikan bahwa suatu dokumen mengandung *term-term* dengan massa yang bervariasi, tidak hanya *term* dengan massa berat dan ringan saja. Dengan kata lain

$$\frac{\sum_{t \in D \cup H(D)} \bar{m}(t)}{|D \cup H(D)|} = c_1 \quad (8)$$

berdasarkan persamaan 6, 7, dan 8 kita dapat menghitung massa dari suatu *term* yang terdapat dalam sebuah dokumen sebagai berikut:

$$m(t, D) = \frac{m(D)\bar{m}(t)}{c_1 \cdot (|D \cup H(D)|)} \quad (9)$$

**Asumsi 4:**

Jumlah *hidden term* dalam sebuah dokumen hanya bergantung pada informasi statistik dari keseluruhan koleksi dokumen dan independen terhadap dokumen itu sendiri.

$$|H(D)| = c_2 \cdot avdl \quad (10)$$

Dimana  $c_2$  adalah konstanta yang bernilai antara 0 sampai  $\sim$ . Karena:

$$|D \cup H(D)| = |D| + |H(D)| \quad (11)$$

Maka

$$|D \cup H(D)| = |D| + c_2 \cdot avdl \quad (12)$$

Selanjutnya

$$|D \cup H(D)| = (1 + c_2) \cdot avdl \cdot \left( (1 - \beta) + \beta \frac{|D|}{avdl} \right) \quad (13)$$

Dimana nilai

$$\beta = \frac{1}{1 + c_2}$$

$|D \cup H(D)| = (1 + c_2) \cdot avdl \cdot \left( (1 - \beta) + \beta \frac{|D|}{avdl} \right)$  untuk penyederhanaan, diasumsikan juga bahwa semua *term* yang terdapat dalam sebuah dokumen memiliki besar

diameter yang sama. Sehingga dari persamaan tersebut kita bisa dapatkan bahwa diameter suatu *term* adalah sebagai berikut:

$$di(t, D) = \frac{di(D)}{|D \cup H(D)|} = \frac{di(D)}{|D| + c_2 \cdot avdl} \quad (14)$$

Dari persamaan 8 dan 14 didapatkan persamaan

$$F_{gbm}(t, D) = \rho(D) \cdot \bar{m}(t) \cdot \frac{c(t,D)}{\frac{1}{\epsilon(D)} \cdot \varphi(D) + c(t,D)} \quad (15)$$

**3. Metodologi Penelitian**

Pada bagian ini dipaparkan tentang metode pengambilan data yang digunakan. Selain itu dibahas metode pengembangan perangkat lunak yang digunakan.

**3.1 Metodologi Pengumpulan Data**

Dalam pelaksanaan penelitian ini penulis menggunakan metode pengumpulan data sebagai berikut :

**Studi Kepustakaan**

Studi kepustakaan ini dimaksudkan untuk pengumpulan dan memperoleh data sekunder dengan cara mempelajari, membaca dan mencatat literatur dari beberapa buku dan sumber yang berkaitan dengan permasalahan di atas.

**Browsing Internet**

Untuk melakukan studi perbandingan dengan sistem-sistem yang telah ada di Internasional maka dilakukan Browsing Internet.

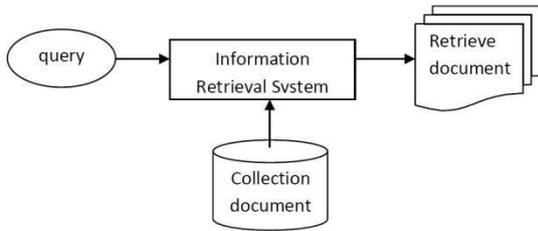
**3.2. Metodologi Pengembangan Perangkat Lunak**

Model proses pengembangan perangkat lunak yang akan digunakan, yaitu model **Linier Sequential / Waterfall**, dimana aktivitas dimulai dengan analisis, perancangan, pengkodean/pemrograman, dan pengujian. Model ini dipergunakan dengan asumsi awal bahwa aplikasi basis data yang disempurnakan/dibangun merupakan sistem yang sudah sangat spesifik dan jelas, sehingga kebutuhan pengguna dapat diidentifikasi secara detail.

**4. Hasil dan Pembahasan**

**4.1. Gambaran Umum Perangkat Lunak**

Perangkat lunak ini akan memberikan hasil berupa *precision* dan *recall* dengan menggunakan model Gravitasi Newton dengan perspektif fisika. Berikut gambaran umum search engine yang akan dibangun :



Gambar 4 Proses dalam Sistem Temu Balik Informasi

Pemodelan Sistem Temu Balik Informasi yang digunakan adalah model ruang vektor dan Model Gravitasi Newton. Prosesnya dibagi dalam dua bagian utama yaitu *Indexing* dan *Searching*. *Indexing* adalah proses yang merepresentasikan koleksi dokumen dalam bentuk indeks sehingga memudahkan dan mempercepat proses pencarian dokumen yang relevan. Sedangkan *searching (matching)* merupakan proses menemukan kembali dokumen yang relevan terhadap *query*.

#### 4.2. Analisis dan Spesifikasi Kebutuhan Perangkat Lunak

Perangkat lunak ini akan mengimplementasikan Sistem Temu Balik Informasi menggunakan model Gravitasi Newton sebagai model *matching* antara query dan dokumen. Perangkat lunak ini bekerja selayaknya search engine dimana perangkat lunak akan menerima query dari *user* dan menghasilkan keluaran berupa kumpulan dokumen yang dianggap paling relevan berdasarkan pemodelan diatas dan selanjutnya akan dibandingkan hasil *precision* dan *recall* dari model Gravitasi Newton versi Diskrit dan model Gravitasi Newton versi Continuous. Perangkat lunak ini diimplementasikan berbasis web.

Masukan untuk perangkat lunak adalah berupa koleksi dokumen. Agar proses pencarian dapat berjalan dengan lancar, koleksi dokumen harus memiliki syarat-syarat sebagai berikut:

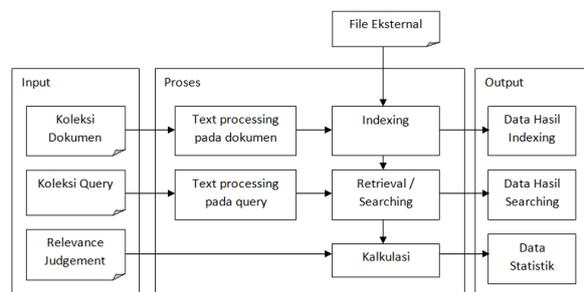
1. Seluruh koleksi dokumen dalam satu format file .txt, untuk memudahkan pembacaan. Dokumen juga harus memiliki id unik untuk memudahkan pemrosesan. Untuk itu dibutuhkan *package* untuk menyimpan kumpulan dari dokumen file teks.
2. Seluruh kata kunci yang akan digunakan sudah ditentukan sebelumnya untuk memudahkan pencarian. Selain koleksi dokumen, perangkat lunak juga memerlukan masukan berupa daftar kata-kata yang termasuk stopword. Daftar stopword ini diperlukan agar saat menghitung kemunculan kata stopword tidak ikut dihitung oleh perangkat lunak.

Proses yang terjadi dalam perangkat lunak terdiri dari *indexing* dan *searching*. Kebutuhan perangkat lunak untuk kedua proses ini adalah:

1. Membaca dokumen file  
Perangkat lunak harus mampu membaca file dokumen sebagai bagian dari proses *indexing*. Selain itu, file yang telah dibaca juga mengalami proses *stopword* dan *stemming*. Selanjutnya, dilakukan proses pembobotan setiap term yang ada pada dokumen tersebut.
2. Membaca query awal inputan pengguna (*user*).  
Perangkat lunak harus mampu baca query inputan user serta melakukan proses *stopword* dan *stemming* pada query tersebut.
3. Menghitung *Similarity* dan Pembobotan.  
Perhitungan *Similarity* merupakan bagian dari proses *searching*. Nilai bobot dokumen dan bobot query yang telah didapatkan pada proses sebelumnya, digunakan untuk dapat nilai *similarity* antara query dan dokumen.
4. Reranking dokumen berdasarkan nilai *Similarity*.  
Perangkat lunak harus mampu mengurutkan dokumen berdasarkan nilai *Similarity* yang paling besar. Dokumen teratas pada pengurutan dianggap paling relevan terhadap query.
5. Menghitung *Precision* dan *Recall* perangkat lunak.  
Perangkat lunak harus mampu menghitung nilai *Precision* dan *Recall* sehingga dapat diketahui performansi dari perangkat lunak.

#### 4.3. Arsitektur Perangkat Lunak

Arsitektur perangkat lunak yang dibangun memiliki input utama berupa koleksi dokumen. Sebelum fungsi lainnya dapat dilakukan, koleksi dokumen sudah harus selesai dikonversi menjadi *inverted file*. Dalam pemrosesannya menjadi *inverted file* tersebut dilakukan *indexing* dan *stemming*. *inverted file* disimpan di dalam basis data. Input berupa *query collection* digunakan untuk melakukan *retrieval* dokumen koleksi. *Relevance judgement* digunakan untuk melakukan penghitungan performansi sistem temu balik informasi berupa *recall* dan *precision*. *Output* berupa data statistik akan ditampilkan berupa data-data *recall* dan *precision*. Gambar 5 di bawah adalah arsitektur perangkat lunak yang dibangun.



Gambar 5 Arsitektur Perangkat Lunak

#### 4.4. Analisis Kebutuhan Fungsional Perangkat Lunak

Kebutuhan fungsional dari perangkat lunak adalah sebagai berikut :

1. Dapat mengolah data koleksi dokumen jadi koleksi database yang berisi kumpulan term yang ada pada koleksi dokumen, seperti informasi jumlah kemunculan suatu term pada dokumen serta nilai pembobotan setiap term pada koleksi dokumen.
2. Dapat mengambil informasi term dari koleksi database dan mengolahnya untuk mendapatkan kumpulan dokumen yang relevan dari koleksi dokumen terhadap query.
3. Dapat melakukan perhitungan similiarity antara query dengan koleksi dokumen yang telah diolah menjadi koleksi database dan kemudian dapat menampilkan hasil perhitungan nilai similiarity yang telah dilakukan.
4. Dapat menghitung performansi dari perangkat lunak berdasarkan parameter uji yaitu *Precision* dan *Recall*.

Untuk menghitung *Precision*

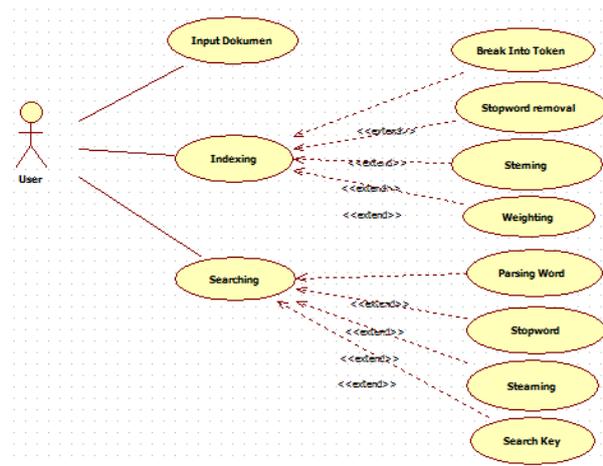
$$precision = \frac{\text{jumlah dokumen relevan yang berhasil ditemukan}}{\text{jumlah dokumen yang dikembalikan sistem}}$$

Untuk menghitung *Recall*

$$recall = \frac{\text{jumlah dokumen relevan yang berhasil ditemukan}}{\text{jumlah dokumen relevan pada koleksi}}$$

4.5. *Pemodelan Analisis*

Pemodelan hasil analisis digambarkan dengan menggunakan use case diagram. Gambar 6 di bawah merupakan use case diagram dari perangkat lunak yang dibuat.



Gambar 6 Use Case Perangkat Lunak

4.6. *Perancangan Modul Perangkat Lunak*

Perangkat lunak terbagi ke dalam 6 modul, yang dapat dilihat di tabel di bawah ini :

Tabel 2 Modul Perangkat Lunak

No	Modul	Deskripsi
1	Indexing	Menerima masukan berupa koleksi dokumen untuk diubah menjadi <i>inverted file</i> dan disimpan ke basis data.
2	Query relevance	Membaca <i>file</i> eksternal dan menyimpan <i>queryrelevance</i> ke basis data
3	Searching	Melakukan <i>retrieval</i> dokumen berdasarkan input koleksi <i>query</i> dan <i>inverted file</i> yang terbentuk lalu mengevaluasi kinerja sistem dengan menampilkan <i>relevance judgement</i> .
4	Database	Melakukan koneksi ke database
5	Fungsi	Kumpulan fungsi untuk mendukung keempat modul lainnya
6	User Interface	Mengatur tampilan program secara keseluruhan seperti margin, layout, warna dll.

4.7. *Perancangan Antar Muka*

a. Rancangan Antar Muka Halaman Indexing

Gambar 7 di bawah menggambarkan antar muka halaman indexing



Gambar7 Antar Muka Halaman Indexing

b. Rancangan Antar Muka Halaman Searching

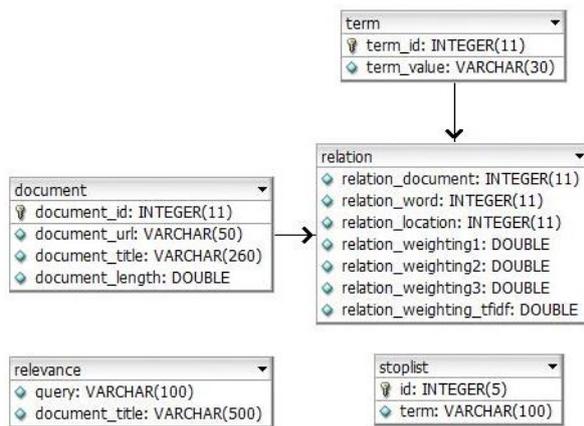
Gambar 8 di bawah menggambarkan antar muka halaman searching.



Gambar 8 Antar Muka Halaman Searching

4.8. *Perancangan Basis Data*

Rancangan basis data perangkat lunak dapat dilihat pada gambar 9 dibawah ini.



Gambar 9 Rancangan Basis Data

## 5. Kesimpulan

Pada bagian ini akan diuraikan hal yang dapat disimpulkan dari proses yang dilakukan. Selain itu diuraikan pula beberapa saran yang dapat digunakan dalam pengembangan selanjutnya..

### 5.1 Simpulan

Berdasarkan hasil analisis dan perancangan perangkat lunak yang dilakukan dapat diambil beberapa kesimpulan, yaitu:

1. Nilai *recall* pada seluruh metode adalah sama karena jumlah dokumen relevan keseluruhan dan dokumen terambil sama untuk ketiga metode dan tidak diperhitungkannya keterurutan dalam dokumen yang *re-trieve*.
2. Metode pembobotan Inverse kuadrat memberikan nilai pembobotan yang lebih baik dibandingkan dengan metode Eksponensial negatif dan metode diksrit dalam *Gravitation Based Model* dapat dilihat dari hasil perhitungan *precision*.

## 5.2 Saran

Untuk pengembangan, disarankan hal-hal sebagai berikut:

1. *Document collection* untuk model Gravitasi Newton menggunakan jenis document lain selain *free text (Document collection support* perhitungan *Pagerank*).
2. Model Gravitasi Newton dapat digunakan sebagai perbandingan kinerja Sistem Temu Balik Informasi dengan model pendekatan yang lain.

## 6. Daftar Rujukan

- [1] Wibowo, Ari. Peningkatan Performansi Sistem Temu Balik Informasi dengan Metode *Phrasal Translation* dan *Query Expansion*. Batam: Politeknik Negeri Batam..
- [2] [https://en.wikipedia.org/wiki/Information\\_retrieval](https://en.wikipedia.org/wiki/Information_retrieval). Diakses pada tanggal 9 September 2017.
- [3] Mandala, Rila. 2002. *Peningkatan Performansi Sistem Temu-Kembali Informasi dengan Perluasan Query Secara Otomatis*. Institut Teknologi Bandung.
- [4] Moffat, Alistair., Justin Zobel and David Hawking. 2004. *Recommended Reading for STBI Research Students*. Australia: The University of Melbourne.
- [5] Dudi Indrajit. 2009. *Mudah dan Aktif Belajar Fisika 3 (IPA)*. Penerbit: Pusat Perbukuan Departemen Pendidikan Nasional.
- [6] S. Shi, J.R. Wen, Q. YU, R. Song, and W.Y. Ma. 2007. *Model Gravitasi Newton for Information Retrieval*. Microsoft Research.