

# BAB IV

## ANALISA DAN PERANCANGAN

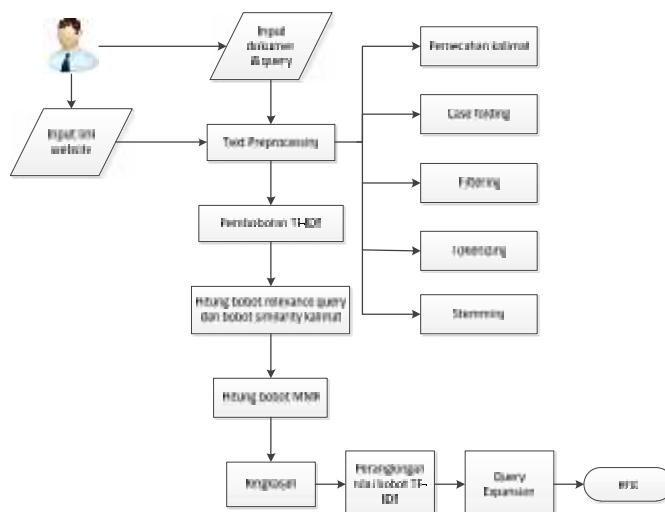
### 4.1 Analisa Sistem Lama

Pada sistem peringkasan dokumen sebelumnya sistem sudah bisa dijalankan namun masih adanya kekurangan pada sistem tersebut yaitu penginputan dokumen yang harus dipindahkan kedalam notepad terlebih dahulu. Disimpan dalam bentuk txt. Itu membuat pengguna menjadi susah.

### 4.2 Analisa Sistem Baru

Dengan kendala yang terdapat pada sistem yang lama tersebut maka dibuatlah suatu sistem baru yang lebih menghemat waktu untuk mengetahui intisari dari dokumen tersebut. Yaitu dengan cara mengcopy dokumen yang akan diringkas ke sistem. Sedangkan dokumen *online* yang ingin diringkas dapat menginputkan *link website* nya saja kedalam sistem. Kemudian membuat *query expansion* untuk mengetahui apakah judul tersebut sudah menggambarkan isi dari ringkasan tersebut. menghitung akurasi sistem.

Berikut ini adalah analisa dan perancangan aplikasi *Automatic Document Summarization*.



Gambar 4.1 Analisa Sistem

Proses secara umum dalam pembuatan ringkasan otomatis pada penelitian ini, yaitu *text preprocessing*, pembobotan TF-IDF, pembobotan *relevance query*, pembobotan *similarity* kalimat dan ekstraksi MMR. Ketika teks akan diringkas proses yang akan dilakukan adalah :

1. *User* memasukkan dokumen dan menuliskan *query* (judul) dari dokumen tersebut.
2. Kemudian sistem melakukan pemrosesan teks (*text preprocessing*), yaitu pemecahan kalimat, *case folding*, *filtering*, *tokenizing* kata, dan *stemming*.
3. Hasil dari pemrosesan teks kemudian dilakukan pembobotan TF-IDF. kemudian menghitung bobot *relevance* antar *query* terhadap seluruh kalimat dalam dokumen dan menghitung bobot *similarity* antar kalimat.
4. Hasil dari pembobotan TF-IDF dan bobot *query relevance* digunakan untuk menghitung bobot MMR kalimat dengan kombinasi matrik bobot *relevance* dan bobot *similarity* kalimat. Pembobotan TF-IDF ini juga digunakan untuk membuat *query expansion*. Yang mana nilai bobot kata dirangking kemudian 5 bobot kata yang tertinggi akan dijadikan *query expansion*.
5. Tahap terakhir yaitu ekstraksi menghitung bobot MMR kalimat.

#### **4.2.1 Data Masukan (*Input*)**

Data masukan (*input*) yang ada pada sistem ini adalah :

1. Dokumen Manual yaitu dokumen yang di inputkan secara manual.
2. Dokumen *Website* yaitu dokumen yang di inputkan dengan cara *mengcopy link website* saja.

#### 4.2.2 Proses

Proses yang terjadi pada sistem ini adalah :

1. Proses ringkasan dokumen manual :  
*User* melakukan operasi *input* dan ubah pada proses ringkasan dokumen manual. Yang terdiri dari : *Judul/Query, Teks* dokumen dan *compression*.
2. Proses ringkasan dokumen *website* :  
*User* melakukan proses *input* dan ubah pada proses ringkasan dokumen *website*. Yang terdiri dari : *Link website* dan *compression*.

#### 4.2.3 Data Keluaran (*output*)

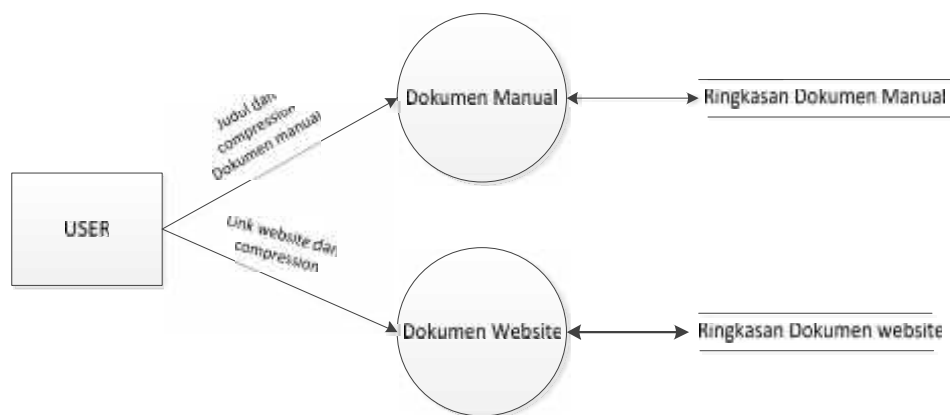
Bentuk keluaran (*output*) yang akan ditampilkan pada sistem ini adalah berupa Ringkasan dokumen.

### 4.3 Deskripsi Fungsional

Deskripsi fungsional adalah gambaran umum sistem secara umum yang akan dirancang.

#### 4.3.1 Data Flow Diagram (DFD)

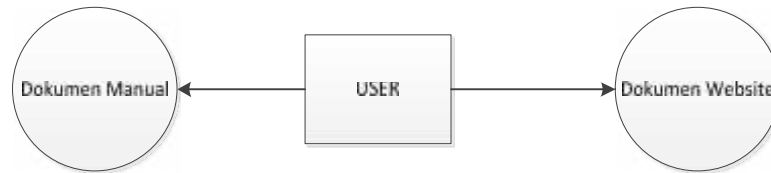
*Data Flow Diagram* (DFD) sering digunakan untuk menggunakan suatu sistem yang telah ada atau sistem baru yang akan dikembangkan secara logika tanpa mempertimbangkan lingkungan fisik dimana data tersebut mengalir, atau lingkungan fisik dimana data tersebut tersimpan.



Gambar 4.2 DFD level 1

### 4.3.2 Entity Relation Diagram (ERD)

ERD (*Entity Relation Diagram*) adalah gambaran mengenai berelasinya antar entitas. Berikut adalah ERD dari sistem peringkas dokumen.



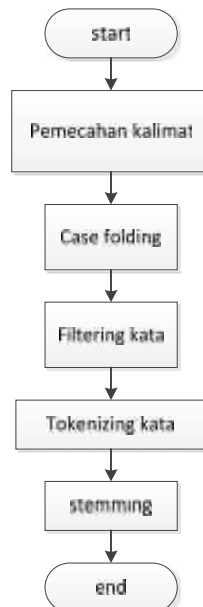
Gambar 4.3 ERD

### 4.3.3 Flowchart

Proses-proses yang terjadi pada sistem peringkas dokumen bisa digambarkan dengan menggunakan *flowchart*

#### 4.3.3.1 Analisa *Text Preprocessing*

pada tahap ini dokumen yang telah diinputkan beserta *query* akan dilakukan proses *text preprocessing*. Berikut ini adalah *flowchart text preprocessing* :



Gambar 4.4 *Flowchart text preprocessing*

Berikut adalah contoh dokumen yang akan diinputkan dalam proses *text preprocessing* :

Judul : "Tidak mau kalah, kepolisian amerika Gunakan Lamborghini Gallardo"

Tidak mau kalah dengan para polisi di daerah timur tengah, kini kepolisian Amerika Serikat (AS) khususnya Los Angeles dapat membusungkan dada mereka. Pasalnya kini Los Angeles Police Department yang biasa dikenal dengan LAPD menambah skuad armada tempurnya dengan menggunakan supercar asal Italia, Lamborghini Gallardo.

Mobil yang dapat melaju dari kecepatan 0 hingga 100 km/jam hanya dalam waktu 3,7 detik ini dibuat antara tahun 2013 hingga 2014. Beruntungnya adalah, LAPD tidak membeli kendaraan tersebut, melainkan hibah dari seorang pemilik perusahaan fiber optic di Amerika, Light Source bernama Nathale Marg dan Travis Marg.

Lamborghini Gallardo... adalah mobil super dengan menggunakan mesin V10 dengan kapasitas 5,0 liter atau mesin V10 5,2 liter. Mesin tersebut dikawinkan dengan transmisi manual 5-kecepatan atau dengan pilihan transmisi 6-kecepatan otomatis.

Kehadiran mobil super LAPD yang diberikan seorang donatur guna mendukung sebuah kegiatan amal yang bermaksud untuk mendukung sebuah solusi untuk trafik LAPD

Dina foto yang terlihat dari akun Twitter dan Instagram, mobil ini diberi nama LAPD number 1 terlihat mobil super ini memiliki desain khas LAPD dengan dominan putih dan sudah dilengkapi dengan sirene serta perlengkapan Light bar, kerya mobil police

Gambar 4.5 Contoh berita

#### 4.3.3.1.1 Analisa Pemecahan Kalimat

Yang dilakukan pada proses ini adalah memecah *string* dokumen utuh menjadi kalimat-kalimat dengan menghilangkan delimiter atau tanda baca yang menyusunnya yaitu seperti titik “.”, tanda tanya “?”, dan tanda seru “!”.



Gambar 4.6 Flowchart pemecahan kalimat

Hasil pemecahan kalimat dari contoh riset pada gambar 4.5 dapat di lihat pada tabel dibawah ini.

Tabel 4.1 Hasil Pemecahan Kalimat

No	Kalimat
1	Tidak mau kalah dengan para polisi di daerah timur tengah, kini kepolisian Amerika Serikat (AS) khususnya Los Angeles dapat membusungkan dada mereka
2	Pasalnya kini Los Angeles Police Department yang biasa dikenal dengan LAPD menambah skuad armada tempurnya dengan menggunakan supercar asal Italia, Lamborghini Gallardo
3	Mobil yang dapat melaju dari kecepatan 0 hingga 100 km/jam hanya dalam waktu 3,7 detik ini dibuat antara tahun 2013 hingga 2014
4	Beruntungnya adalah, LAPD tidak membeli kendaraan tersebut, melainkan hibah dari seorang pemilik perusahaan fiber optic di Amerika, Light Source bernama Nathale Marg dan Travis Marg
5	Lamborghini Gallardo adalah mobil super dengan menggunakan mesin V10 dengan kapasitas 5,0 liter atau mesin V10 5,2 liter
6	Mesin tersebut dikawinkan dengan transmisi manual 6-kecepatan atau dengan pilihan transmisi 6-kecepatan otomatis
7	Kehadiran mobil super LAPD yang diberikan seorang donatur guna mendukung sebuah kegiatan amal yang bermaksud untuk mendukung sebuah satuan udara milik LAPD.
8	Dari foto yang terlihat dari akun Twitter dan Instagram, mobil ini diberi nama LAPDLambo
9	Terlihat mobil super ini memiliki desain khas LAPD dengan dominan putih dan sudah dilengkapi dengan sirene serta kelengkapan lain layaknya mobil polisi

#### 4.3.3.1.2 Analisa Case Folding

Dokumen yang telah dipotong menjadi kalimat kemudian pada tahap ini dicase folding, proses yang dilakukan yaitu mengubah teks menjadi huruf kecil, menghilangkan angka dan tanda baca maupun simbol-simbol karena dianggap sebagai delimiter, sistem hanya menerima karakter huruf saja.



Gambar 4.7 *Flowchart Case Folding*

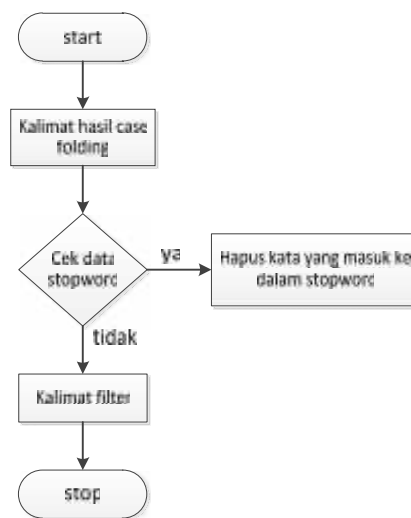
Hasil *Case folding* dari contoh riset pada gambar 4.5 dapat di lihat pada tabel 4.2.

Tabel 4.2 Hasil *Case Folding*

No	Kalimat
1	tidak mau kalah dengan para polisi di daerah timur tengah kini kepolisian amerika serikat as khususnya los angeles dapat membusungkan dada mereka
2	pasalnya kini los angeles police department yang biasa dikenal dengan lapd menambah skuad armada tempurnya dengan menggunakan supercar asal Italia lamborghini gallardo.
3	mobil yang dapat melaju dari kecepatan hingga km jam hanya dalam waktu detik ini dibuat antara tahun hingga
4	beruntungnya adalah lapd tidak membeli kendaraan tersebut melainkan hibah dari seorang pemilik perusahaan fiber optic di amerika light source bernama nathale marg dan travis marg
5	lamborghini gallardo adalah mobil super dengan menggunakan mesin v dengan kapasitas liter atau mesin v liter
6	mesin tersebut dikawinkan dengan transmisi manual kecepatan atau dengan pilihan transmisi kecepatan otomatis
7	kehadiran mobil super lapd yang diberikan seorang donatur guna mendukung sebuah kegiatan amal yang bermaksud untuk mendukung sebuah satuan udara milik lapd
8	dari foto yang terlihat dari akun twitter dan instagram mobil ini diberi nama lapdlambo
9	terlihat mobil super ini memiliki desain khas lapd dengan dominan putih dan sudah dilengkapi dengan sirene serta kelengkapan lain layaknya mobil polisi

#### 4.3.3.1.3 Filtering

pada tahap ini yang dilakukan adalah tahap mengambil kata-kata penting dari hasil *case folding* kalimat dan membuang kata-kata yang dianggap kurang penting. Algoritma yang dipakai yaitu *stopword* (membuang kata yang kurang penting). Dalam tahap pembuangan kata-kata yang tidak penting adalah kata yang hasil parsing dicek dengan kamus *stopword* (kumpulan kata). Jika kata parsing ada yang sama dengan *stopword* maka kata akan dibuang atau dihapus. Daftar *stopword* yang digunakan dalam penelitian ini di ambil dari penelitian Tala (2003) dapat dilihat pada halaman lampiran.



Gambar 4.8 Flowchart Filtering Kata

Hasil *Filtering* kata dari contoh riset pada gambar 4.5 dapat di lihat pada tabel dibawah ini.

Tabel 4.3 Hasil *Fitering* kata

No.	Kalimat
1	kalah polisi daerah timur tengah kepolisian amerika serikat as khususnya los angeles membusungkan dada
2	pasalnya los angeles police department dikenal lapd menambah skuad armada tempurnya supercar Italia lamborghini gallardo.

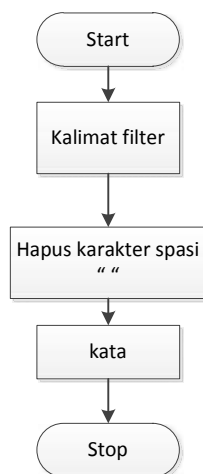


Tabel 4.3 Hasil *Fitering* kata (lanjutan)

No.	Kalimat
3	mobil melaju kecepatan km jam detik
4	beruntungnya lapd membeli kendaraan melainkan hibah pemilik perusahaan fiber optic amerika light source bernama nathale marg travis marg
5	lamborghini gallardo mobil super menggunakan mesin v kapasitas liter mesin v liter
6	mesin dikawinkan transmisi manual kecepatan pilihan transmisi kecepatan otomatis
7	kehadiran mobil super lapd diberikan donatur mendukung kegiatan amal bermaksud mendukung satuan udara lapd
8	foto terlihat mobil diberi lapdlambo
9	terlihat mobil super lapd dilengkapi kelengkapan mobil polisi

#### 4.3.3.1.4 *Tokenizing*

Pada tahap ini proses yang dilakukan yaitu pemotongan string kalimat-kalimat hasil *filtering* berdasarkan delimiter yang menyusunnya yaitu karakter spasi (“ ”).



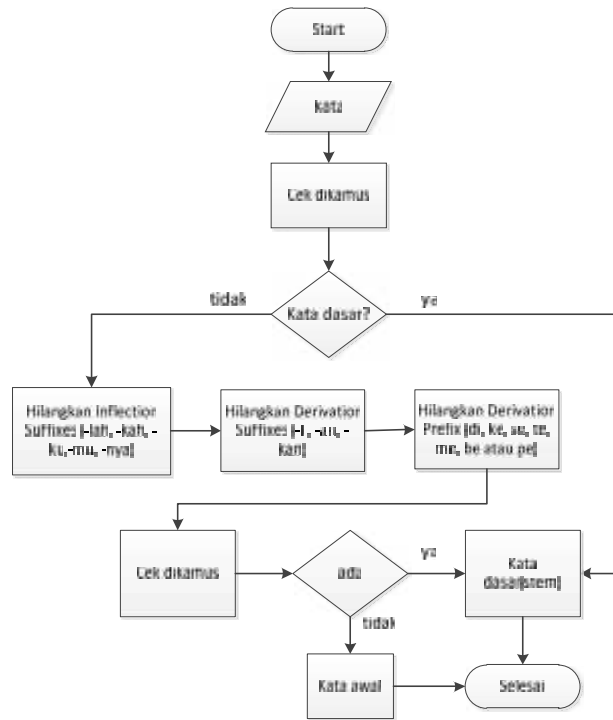
Gambar 4.9 *Flowchart Tokenizing*

Tabel 4.4 Hasil *Tokenizing* kata

Kata	Kata	Kata	Kata
kalah	mobil	dikawinkan	kegiatan
polisi	melaju	transmisi	amal
daerah	kecepatan	manual	bermaksud
timur	km	pilihan	satuan
tengah	jam	otomatis	udara
super	detik	kehadiran	gallardo
amerika	beruntungnya	diberikan	liter
serikat	membeli	donatur	kapasitas
as	kendaraan	mendukung	lamborghini
khususnya	melainkan	pasalnya	mesin
los	hibah	police	light
angeles	pemilik	departement	source
membusungkan	perusahaan	dikenal	v
dada	fiber	bernama	-
optic	lapd	nathale	-

#### 4.3.3.1.5 *Stemming* dengan Algoritma Nazief dan Andriani

Pada bagian ini setelah melakukan *tokenizing* setiap kata yang ada pada dokumen akan dilakukan *stemming* yang bertujuan untuk mendapatkan kata dasar dari setiap kata yang ada pada dokumen. Pada penelitian ini menggunakan metode *stemming Nazief dan Andriani*. Seperti yang terlihat pada gambar 4.10 *flowchart* dari *stemming Nazief dan Andriani*. Hasil *Stemming* kata dari contoh riset pada gambar 4.5 dapat di lihat pada tabel 4.5.



Gambar 4.10 *Flowchart Stemming* Nazief dan Andriani

Tabel 4.5 Hasil *Stemming*

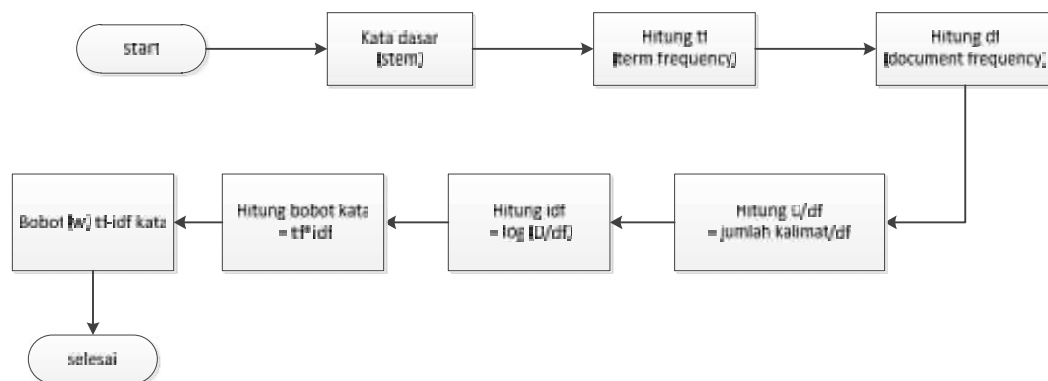
Kata	Kata	Kata	Kata
armada	fiber	lapdlambo	polisi
akun	foto	layak	putih
amal	Gallardo	lengkap	satu
amerika	guna	light	serikat
angeles	hadir	liter	sirene
as	hibah	los	source
beli	instagram	maksud	skuat
beri	italia	manual	super
berik	jam	marg	supercar
buat	kalah	laju	tambah
busung	kapasitas	mesin	tempur

Tabel 4.5 Hasil *Stemming* (Lanjutan)

Kata	Kata	Kata	Kata
cepat	kawin	milik	timur
dada	kenal	mobil	transmisi
daerah	kendara	nama	travis
departement	khas	nathale	twitter
desain	khusus	optic	udara
detik	km	otomatis	untung
dominan	lain	pasal	usaha
donatur	lamborghini	pilih	v
dukung	lapd	police	-

#### 4.3.3.2 Analisa Pembobotan TF-IDF

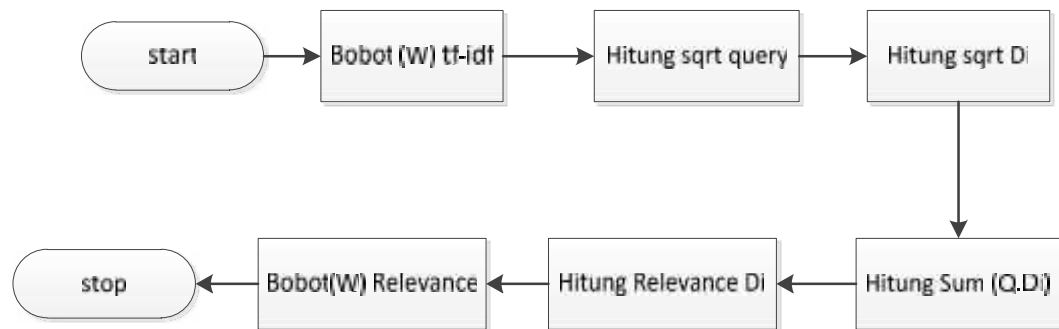
Pembobotan secara otomatis biasanya berdasarkan jumlah kemunculan suatu kata dalam sebuah dokumen (*term frequency*) dan jumlah kemunculannya dalam koleksi dokumen (*inverse document frequency*). Bobot kata semakin besar jika sering muncul dalam suatu dokumen dan semakin kecil jika muncul dalam banyak dokumen. Pembobotan tf-idf dilakukan untuk mengetahui kata apa yang merupakan *query expansion* yang mana bobot akan dirangking kemudian bobot yang tertinggi akan dijadikan *query expansion*. Kemudian pembobotan tf-idf ini juga dilakukan untuk tahap selanjutnya yaitu digunakan untuk mengetahui bobot *relevance query* dan bobot *similarity* kalimat. Perhitungan bobot tf-idf dapat dilihat pada lampiran. Berikut ini adalah flowchart pembobotan tf-idf.



Gambar 4.11 *Flowchart* Pembobotan TF-IDF

#### 4.3.3.3 Analisa Pembobotan *Relevance Query*

Perhitungan bobot *query relevance* merupakan bobot hasil perbandingan kemiripan (similaritas) antara *query* yang dimasukkan oleh *user* terhadap keseluruhan kalimat. Perhitungan bobot *query relevance* ini menggunakan metode *cosine similarity* dengan menghitung cosinus sudut dari dua vector, yaitu  $W$ , bobot dari tiap kalimat dan  $W$  (bobot) *query*. Berikut adalah *flowchart* Pembobotan *Relevance Query*.



Gambar 4.12 *Flowchart* Pembobotan *Relevance Query*

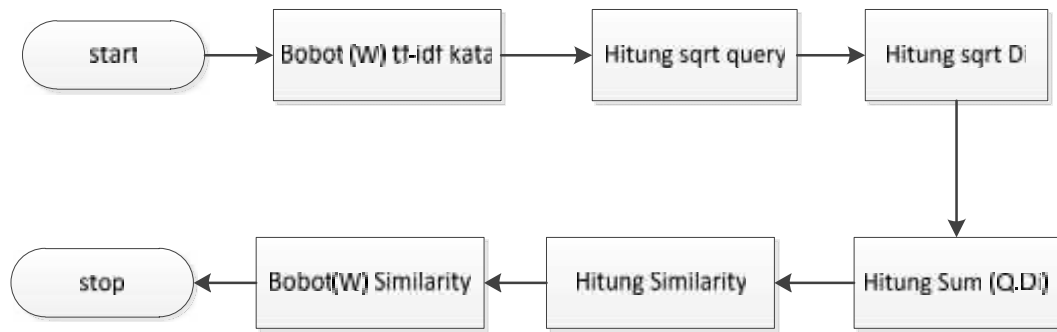
Berdasarkan perhitungan bobot *relevance query* yang telah dilakukan. Maka berikut adalah hasil bobot *relevance query* tiap kalimat atau tiap dokumen. Perhitungan bobot *Relevance Query* secara lengkap dapat dilihat pada lampiran.

Tabel 4.6 Hasil Bobot *Relevance Query*

D1	D2	D3	D4	D5	D6	D7	D8	D9
0.364	0.137	0	0.055	0.137	0	0.142	0	0.071

#### 4.3.3.4 Analisa Pembobotan *Similarity* Kalimat

bobot *similarity* kalimat merupakan bobot hasil perbandingan kemiripan antar kalimat. Berikut adalah *flowchart* Pembobotan *Similarity* Kalimat.



Gambar 4. 13 Flowchart Pembobotan *Similarity* Kalimat

Tabel 4.7 Hasil Bobot *Similarity* antar kalimat

	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>	<b>D5</b>	<b>D6</b>	<b>D7</b>	<b>D8</b>	<b>D9</b>
<b>D1</b>	1	0.081	0	0.032	0	0	0	0	0.084
<b>D2</b>	0.081	1	0	0.009	0.078	0	0.022	0	0.012
<b>D3</b>	0	0	1	0	0.009	0.124	0.008	0.012	0.018
<b>D4</b>	0.032	0.009	0	1	0	0	0.018	0	0.009
<b>D5</b>	0	0.078	0.009	0	1	0.084	0.026	0.008	0.034
<b>D6</b>	0	0	0.124	0	0.084	1	0	0	0
<b>D7</b>	0	0.022	0.008	0.018	0.026	0	1	0.008	0.056
<b>D8</b>	0	0	0.012	0	0.008	0	0.008	1	0.017
<b>D9</b>	0.084	0.012	0.018	0.009	0.034	0	0.056	0.017	1

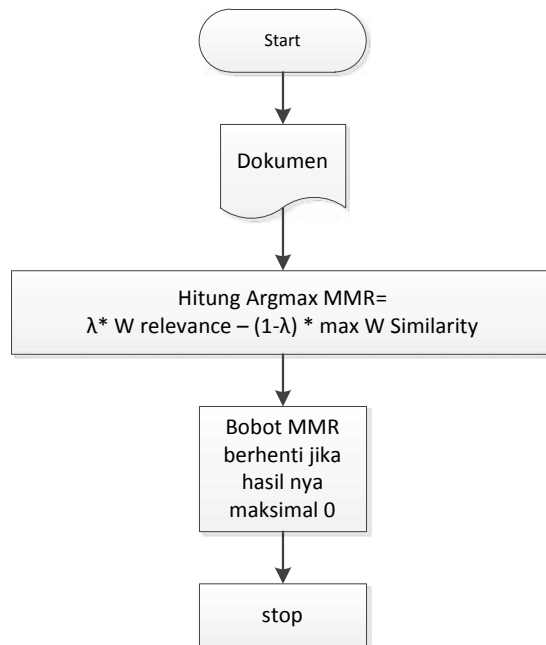
Berikut adalah Hasil Bobot similarity antar kalimat pada tabel 4.7. Hasil ini didapat dari perhitungan bobot *similarity* antar kalimat. Hasil ini digunakan untuk perhitungan selanjutnya. Perhitungan bobot *Similarity* kalimat secara lengkap dapat dilihat pada lampiran.

#### 4.3.3.5 Analisa *Maximum Marginal Relevance*

Algoritma *Maximum Marginal Relevance* (MMR) merupakan salah satu metode ekstraksi ringkasan (*extractive summary*) yang digunakan untuk meringkas dokumen tunggal atau multi dokumen. MMR meringkas dokumen dengan menghitung kesamaan (*similarity*) antara bagian teks. MMR digunakan

untuk merangking kalimat-kalimat sebagai tanggapan terhadap *query* yang diberikan oleh *user*. Perhitungan MMR dilakukan dengan iterasi dengan mengkombinasikan dua matrik *cosine similarity* yaitu relevansi antara *query* terhadap keseluruhan kalimat dan *similarity* antara kalimat dengan kalimat.

Kalimat dengan nilai MMR tertinggi dari setiap perhitungan iterasi akan diambil, kemudian dipilih untuk dijadikan sebagai ringkasan. Iterasi akan berhenti ketika hasil MMR maksimum sama dengan 0 atau 1. Untuk peringkasan *small dokumen*, seperti pada berita (*news*), menggunakan nilai parameter  $\lambda=0.7$ , karena akan menghasilkan ringkasan yang baik (Jade Goldstein, 2008). Berikut adalah *flowchart Maximum Marginal Relevance*.



Gambar 4.14 *Flowchart MMR*

Dari hasil perhitungan bobot *relevance query* dan *similarity* kalimat, tahap selanjutnya kedua matrik ini dikombinasikan pada perhitungan bobot MMR.

$$MMR = \operatorname{argmax} [\lambda * Sim_1(S_i, Q) - (1 - \lambda) * \max Sim_2(S_i, S')]$$

Adapun nilai parameter  $\lambda$  yang digunakan pada perhitungan MMR adalah  $\lambda=0.7$ .  $Sim_1(S_i, Q)$  adalah *relevance query*. Sedangkan  $Sim_2(S_i, S')$  adalah *similarity* kalimat terhadap kalimat yang diekstrak.

1. Perhitungan iterasi ke-1 MMR

Pada iterasi MMR pertama nilai  $Sim_2(S_i, S')$  adalah 0 atau kosong.

$$\begin{aligned}MMR(1) &= [0.7 * Sim_1(D_1, Q) - (1 - 0.7) * maxSim_2(D_1, D')] \\ &= [0.7 * 0.364 - (0.3) * 0] = 0.255\end{aligned}$$

$$\begin{aligned}MMR(2) &= [0.7 * Sim_1(D_2, Q) - (1 - 0.7) * maxSim_2(D_2, D')] \\ &= [0.7 * 0.137 - (0.3) * 0] = 0.096\end{aligned}$$

$$\begin{aligned}MMR(3) &= [0.7 * Sim_1(D_3, Q) - (1 - 0.7) * maxSim_2(D_3, D')] \\ &= [0.7 * 0 - (0.3) * 0] = 0\end{aligned}$$

$$\begin{aligned}MMR(4) &= [0.7 * Sim_1(D_4, Q) - (1 - 0.7) * maxSim_2(D_4, D')] \\ &= [0.7 * 0.055 - (0.3) * 0] = 0.039\end{aligned}$$

$$\begin{aligned}MMR(5) &= [0.7 * Sim_1(D_5, Q) - (1 - 0.7) * maxSim_2(D_5, D')] \\ &= [0.7 * 0.137 - (0.3) * 0] = 0.096\end{aligned}$$

$$\begin{aligned}MMR(6) &= [0.7 * Sim_1(D_6, Q) - (1 - 0.7) * maxSim_2(D_6, D')] \\ &= [0.7 * 0 - (0.3) * 0] = 0\end{aligned}$$

$$\begin{aligned}MMR(7) &= [0.7 * Sim_1(D_7, Q) - (1 - 0.7) * maxSim_2(D_7, D')] \\ &= [0.7 * 0.142 - (0.3) * 0] = 0.099\end{aligned}$$

$$\begin{aligned}MMR(8) &= [0.7 * Sim_1(D_8, Q) - (1 - 0.7) * maxSim_2(D_8, D')] \\ &= [0.7 * 0 - (0.3) * 0] = 0\end{aligned}$$

$$\begin{aligned}MMR(9) &= [0.7 * Sim_1(D_9, Q) - (1 - 0.7) * maxSim_2(D_9, D')] \\ &= [0.7 * 0.071 - (0.3) * 0] = 0.050\end{aligned}$$



Dari hasil perhitungan iterasi ke-1 bobot  $\text{ArgmaxMMR} = 0.255$  pada  $D_1$  atau kalimat ke-1, maka dipilih sebagai kalimat yang akan dijadikan kalimat ringkasan.

## 2. Perhitungan iterasi ke-2 MMR

Pada iterasi ke-2 hasil argmax kalimat pada iterasi ke-1 kemudian digunakan untuk menghitung *similarity* pada  $\max \text{Sim}_2(S_i, S')$  yaitu  $\text{Sim}_2(D_i, D_1)$ .

$$\begin{aligned} \text{MMR}(2) &= [0.7 * \text{Sim}_1(D_2, Q) - (1 - 0.7) * \max \text{Sim}_2(D_2, D_1)] \\ &= [0.7 * 0.137 - (0.3) * 0.081] = 0.072 \end{aligned}$$

$$\begin{aligned} \text{MMR}(3) &= [0.7 * \text{Sim}_1(D_3, Q) - (1 - 0.7) * \max \text{Sim}_2(D_3, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\begin{aligned} \text{MMR}(4) &= [0.7 * \text{Sim}_1(D_4, Q) - (1 - 0.7) * \max \text{Sim}_2(D_4, D_1)] \\ &= [0.7 * 0.055 - (0.3) * 0.032] = 0.029 \end{aligned}$$

$$\begin{aligned} \text{MMR}(5) &= [0.7 * \text{Sim}_1(D_5, Q) - (1 - 0.7) * \max \text{Sim}_2(D_5, D_1)] \\ &= [0.7 * 0.137 - (0.3) * 0] = 0.096 \end{aligned}$$

$$\begin{aligned} \text{MMR}(6) &= [0.7 * \text{Sim}_1(D_6, Q) - (1 - 0.7) * \max \text{Sim}_2(D_6, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\begin{aligned} \text{MMR}(7) &= [0.7 * \text{Sim}_1(D_7, Q) - (1 - 0.7) * \max \text{Sim}_2(D_7, D_1)] \\ &= [0.7 * 0.142 - (0.3) * 0] = 0.099 \end{aligned}$$

$$\begin{aligned} \text{MMR}(8) &= [0.7 * \text{Sim}_1(D_8, Q) - (1 - 0.7) * \max \text{Sim}_2(D_8, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\begin{aligned} \text{MMR}(9) &= [0.7 * \text{Sim}_1(D_9, Q) - (1 - 0.7) * \max \text{Sim}_2(D_9, D_1)] \\ &= [0.7 * 0.071 - (0.3) * 0.084] = 0.025 \end{aligned}$$

Dari hasil perhitungan iterasi ke-2 bobot  $\text{ArgmaxMMR} = 0.099$  pada D7 atau kalimat ke-7, maka kalimat akan dipilih sebagai kalimat ringkasan.

### 3. Perhitungan iterasi ke-3 MMR

Pada iterasi ke-3 ini untuk menghitung *similarity* maksimum  $\max Sim_2(S_i, S')$ , dicari dengan membandingkan nilai maksimum *similarity* antara D1 dengan D7 berdasarkan hasil argmax kalimat pada iterasi sebelumnya. Dengan cara mencari terlebih dahulu *similarity* maksimum antara D1(kalimat ke-1) terhadap dokumen dan *similarity* maksimum D7(kalimat ke-7) terhadap dokumen. Kemudian dibandingkan antara kedua *similarity* maksimum dicari nilai maksimumnya dan digunakan untuk menghitung  $\max Sim_2(S_i, S')$ . Berikut perbandingan perhitungannya.

Perbandingan nilai *similarity* maksimum (D1, Di) dan *similarity* maksimum (D7,Di).

*Similarity* maksimum (D1,Di) :

$$sim(D_1, D_2) = 0.081$$

$$sim(D_1, D_3) = 0$$

$$sim(D_1, D_4) = 0.032$$

$$sim(D_1, D_5) = 0$$

$$sim(D_1, D_6) = 0$$

$$sim(D_1, D_8) = 0$$

$$sim(D_1, D_9) = 0.084$$

*Similarity* maksimum (D7,Di) :

$$sim(D_7, D_2) = 0.022$$

$$sim(D_7, D_3) = 0.008$$

$$\text{sim}(D_7, D_4) = 0.018$$

$$\text{sim}(D_7, D_5) = 0.026$$

$$\text{sim}(D_7, D_6) = 0$$

$$\text{sim}(D_7, D_8) = 0.008$$

$$\text{sim}(D_7, D_9) = 0.056$$

Jadi *similarity* maksimum antara  $\text{sim}(D_1, D_i)$  dan  $\text{sim}(D_7, D_i)$  adalah 0.084 pada D1.

$$\begin{aligned} \text{MMR}(2) &= [0.7 * \text{Sim}_1(D_2, Q) - (1 - 0.7) * \text{maxSim}_2(D_2, D_1)] \\ &= [0.7 * 0.137 - (0.3) * 0.081] = 0.072 \end{aligned}$$

$$\begin{aligned} \text{MMR}(3) &= [0.7 * \text{Sim}_1(D_3, Q) - (1 - 0.7) * \text{maxSim}_2(D_3, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\begin{aligned} \text{MMR}(4) &= [0.7 * \text{Sim}_1(D_4, Q) - (1 - 0.7) * \text{maxSim}_2(D_4, D_1)] \\ &= [0.7 * 0.055 - (0.3) * 0.032] = 0.029 \end{aligned}$$

$$\begin{aligned} \text{MMR}(5) &= [0.7 * \text{Sim}_1(D_5, Q) - (1 - 0.7) * \text{maxSim}_2(D_5, D_1)] \\ &= [0.7 * 0.137 - (0.3) * 0] = 0.096 \end{aligned}$$

$$\begin{aligned} \text{MMR}(6) &= [0.7 * \text{Sim}_1(D_6, Q) - (1 - 0.7) * \text{maxSim}_2(D_6, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\begin{aligned} \text{MMR}(8) &= [0.7 * \text{Sim}_1(D_8, Q) - (1 - 0.7) * \text{maxSim}_2(D_8, D_1)] \\ &= [0.7 * 0 - (0.3) * 0] = 0 \end{aligned}$$

$$\text{MMR}(9) = [0.7 * \text{Sim}_1(D_9, Q) - (1 - 0.7) * \text{maxSim}_2(D_9, D_1)]$$

$$= [0.7 * 0.071 - (0.3) * 0.084] = 0.025$$

Dari hasil perhitungan iterasi ke-3 bobot  $\text{ArgmaxMMR} = 0.096$  pada D5 atau kalimat ke-5, kemudian dipilih sebagai kalimat yang akan dijadikan kalimat ringkasan.

#### 4. Perhitungan iterasi ke-4 MMR

Pada iterasi keempat ini untuk menghitung *similarity* maksimum  $\max \text{Sim}_2(S_i, S')$ , dicari dengan membandingkan nilai maksimum *similarity* antara D1, D7 dan D5 berdasarkan hasil argmax kalimat pada iterasi sebelumnya. Dengan mencari terlebih dahulu *similarity* maksimum antara D1(kalimat ke-1) terhadap dokumen, D7 (kalimat ke-7) terhadap dokumen, dan D5(kalimat ke-5) terhadap dokumen. Kemudian setelah dapat hasilnya dibandingkan dan dicari nilai maksimumnya dan digunakan untuk menghitung  $\max \text{Sim}_2(S_i, S')$ . Berikut ini perbandingan perhitungannya.

*Similarity* maksimum (D1,Di) :

$$\text{sim}(D_1, D_2) = 0.081$$

$$\text{sim}(D_1, D_3) = 0$$

$$\text{sim}(D_1, D_4) = 0.032$$

$$\text{sim}(D_1, D_5) = 0$$

$$\text{sim}(D_1, D_6) = 0$$

$$\text{sim}(D_1, D_9) = 0.084$$

*Similarity* maksimum (D7,Di) :

$$\text{sim}(D_7, D_2) = 0.022$$

$$\text{sim}(D_7, D_3) = 0.008$$

$$\text{sim}(D_7, D_4) = 0.018$$

$$\text{sim}(D_7, D_6) = 0$$

$$\text{sim}(D_7, D_8) = 0.008$$

$$\text{sim}(D_7, D_9) = 0.056$$

Similarity maksimum (D5,Di) :

$$\text{sim}(D_5, D_2) = 0.078$$

$$\text{sim}(D_5, D_3) = 0.009$$

$$\text{sim}(D_5, D_4) = 0$$

$$\text{sim}(D_5, D_6) = 0.084$$

$$\text{sim}(D_5, D_8) = 0.008$$

$$\text{sim}(D_5, D_9) = 0.034$$

Jadi *similarity* maksimum antara  $\text{sim}(D_1, D_i)$ ,  $\text{sim}(D_7, D_i)$ , dan  $\text{sim}(D_5, D_i)$  adalah 0.084 pada D5.

$$\begin{aligned} \text{MMR}(2) &= [0.7 * \text{Sim}_1(D_2, Q) - (1 - 0.7) * \text{maxSim}_2(D_2, D_5)] \\ &= [0.7 * 0.137 - (0.3) * 0.078] = 0.073 \end{aligned}$$

$$\begin{aligned} \text{MMR}(3) &= [0.7 * \text{Sim}_1(D_3, Q) - (1 - 0.7) * \text{maxSim}_2(D_3, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.009] = -0.003 \end{aligned}$$

$$\begin{aligned} \text{MMR}(4) &= [0.7 * \text{Sim}_1(D_4, Q) - (1 - 0.7) * \text{maxSim}_2(D_4, D_5)] \\ &= [0.7 * 0.055 - (0.3) * 0] = 0.039 \end{aligned}$$

$$\begin{aligned} \text{MMR}(6) &= [0.7 * \text{Sim}_1(D_6, Q) - (1 - 0.7) * \text{maxSim}_2(D_6, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.084] = -0.025 \end{aligned}$$

$$\begin{aligned} MMR(8) &= [0.7 * Sim_1(D_8, Q) - (1 - 0.7) * maxSim_2(D_8, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.008] = -0.003 \end{aligned}$$

$$\begin{aligned} MMR(9) &= [0.7 * Sim_1(D_9, Q) - (1 - 0.7) * maxSim_2(D_9, D_5)] \\ &= [0.7 * 0.071 - (0.3) * 0.034] = 0.040 \end{aligned}$$

Dari hasil perhitungan iterasi ke-4 bobot  $ArgmaxMMR = 0.073$  pada D2 atau kalimat ke-2, kemudian dipilih sebagai kalimat yang akan dijadikan kalimat ringkasan.

#### 5. Perhitungan iterasi ke-5

Pada iterasi kelima ini untuk menghitung *similarity* maksimum  $\max Sim_2(S_i, S')$ , dicari dengan membandingkan nilai maksimum *similarity* antara D1, D7, D5 dan D2 berdasarkan hasil argmax kalimat pada iterasi sebelumnya. Dengan cara mencari terlebih dahulu *similarity* maksimum antara D1 (kalimat ke-1) terhadap dokumen, D7 (kalimat ke-7) terhadap dokumen, D5 (kalimat ke-5), dan D2 (kalimat ke-2). Kemudian dibandingkan dan dicari nilai maksimumnya dan digunakan untuk menghitung  $\max Sim_2(S_i, S')$ .

Berikut ini adalah perbandingan perhitungannya :

*Similarity* maksimum (D1,Di) :

$$sim(D_1, D_3) = 0$$

$$sim(D_1, D_4) = 0.032$$

$$sim(D_1, D_6) = 0$$

$$sim(D_1, D_8) = 0$$

$$sim(D_1, D_9) = 0.08$$

*Similarity* maksimum (D7,Di) :

$$sim(D_7, D_3) = 0.008$$

$$\text{sim}(D_7, D_4) = 0.018$$

$$\text{sim}(D_7, D_6) = 0$$

$$\text{sim}(D_7, D_8) = 0.008$$

$$\text{sim}(D_7, D_9) = 0.056$$

Similarity maksimum (D5,Di) :

$$\text{sim}(D_5, D_3) = 0.009$$

$$\text{sim}(D_5, D_4) = 0$$

$$\text{sim}(D_5, D_6) = 0.084$$

$$\text{sim}(D_5, D_8) = 0.008$$

$$\text{sim}(D_5, D_9) = 0.034$$

Similarity maksimum (D2,Di) :

$$\text{sim}(D_2, D_3) = 0$$

$$\text{sim}(D_2, D_4) = 0.009$$

$$\text{sim}(D_2, D_6) = 0$$

$$\text{sim}(D_2, D_8) = 0$$

$$\text{sim}(D_2, D_9) = 0.012$$

Jadi *similarity* maksimum antara  $\text{sim}(D_1, D_i)$ ,  $\text{sim}(D_7, D_i)$ ,  $\text{sim}(D_5, D_i)$  dan  $\text{sim}(D_2, D_i)$  adalah 0.084 pada D5.

$$\begin{aligned} \text{MMR}(3) &= [0.7 * \text{Sim}_1(D_3, Q) - (1 - 0.7) * \text{maxSim}_2(D_3, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.009] = -0.003 \end{aligned}$$

$$\begin{aligned} MMR(4) &= [0.7 * Sim_1(D_4, Q) - (1 - 0.7) * maxSim_2(D_4, D_5)] \\ &= [0.7 * 0.055 - (0.3) * 0] = 0.039 \end{aligned}$$

$$\begin{aligned} MMR(6) &= [0.7 * Sim_1(D_6, Q) - (1 - 0.7) * maxSim_2(D_6, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.084] = -0.025 \end{aligned}$$

$$\begin{aligned} MMR(8) &= [0.7 * Sim_1(D_8, Q) - (1 - 0.7) * maxSim_2(D_8, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.008] = -0.003 \end{aligned}$$

$$\begin{aligned} MMR(9) &= [0.7 * Sim_1(D_9, Q) - (1 - 0.7) * maxSim_2(D_9, D_5)] \\ &= [0.7 * 0.071 - (0.3) * 0.034] = 0.040 \end{aligned}$$

Dari hasil perhitungan iterasi ke-5 bobot  $ArgmaxMMR = 0.040$  pada D9 atau kalimat ke-9, kemudian dipilih sebagai kalimat yang akan dijadikan kalimat ringkasan.

#### 6. Perhitungan iterasi ke-6

Pada iterasi ke-6 ini untuk menghitung *similarity* maksimum  $\max Sim_2(S_i, S')$ , dicari dengan membandingkan nilai maksimum *similarity* antara D1, D7, D5, D2 dan D9 berdasarkan hasil argmax kalimat pada iterasi sebelumnya. Dengan cara mencari terlebih dahulu *similarity* maksimum antara D1 (kalimat ke-1) terhadap dokumen, D7 (kalimat ke-7) terhadap dokumen, D5 (kalimat ke-5), D2 (kalimat ke-2) dan D9 (kalimat ke-9). Kemudian dibandingkan dan dicari nilai maksimumnya dan digunakan untuk menghitung  $\max Sim_2(S_i, S')$ . Berikut ini adalah perbandingan perhitungannya :

*Similarity* maksimum (D1,Di) :

$$sim(D_1, D_3) = 0$$

$$sim(D_1, D_4) = 0.032$$

$$sim(D_1, D_6) = 0$$



$$\text{sim}(D_1, D_8) = 0$$

Similarity maksimum (D7,Di) :

$$\text{sim}(D_7, D_3) = 0.008$$

$$\text{sim}(D_7, D_4) = 0.018$$

$$\text{sim}(D_7, D_6) = 0$$

$$\text{sim}(D_7, D_8) = 0.008$$

Similarity maksimum (D5,Di) :

$$\text{sim}(D_5, D_3) = 0.009$$

$$\text{sim}(D_5, D_4) = 0$$

$$\text{sim}(D_5, D_6) = 0.084$$

$$\text{sim}(D_5, D_8) = 0.008$$

Similarity maksimum (D2,Di) :

$$\text{sim}(D_2, D_3) = 0$$

$$\text{sim}(D_2, D_4) = 0.009$$

$$\text{sim}(D_2, D_6) = 0$$

$$\text{sim}(D_2, D_8) = 0$$

Similarity maksimum (D9,Di) :

$$\text{sim}(D_9, D_3) = 0.018$$

$$\text{sim}(D_9, D_4) = 0.009$$

$$\text{sim}(D_9, D_6) = 0$$

$$\text{sim}(D_9, D_8) = 0.017$$

Jadi *similarity* maksimum antara  $\text{sim}(D_1, D_i)$ ,  $\text{sim}(D_7, D_i)$ ,  $\text{sim}(D_5, D_i)$ ,  $\text{sim}(D_2, D_i)$  dan  $\text{sim}(D_9, D_i)$  adalah 0.084 pada D5.

$$\begin{aligned} \text{MMR}(3) &= [0.7 * \text{Sim}_1(D_3, Q) - (1 - 0.7) * \text{maxSim}_2(D_3, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.009] = -0.003 \end{aligned}$$

$$\begin{aligned} \text{MMR}(4) &= [0.7 * \text{Sim}_1(D_4, Q) - (1 - 0.7) * \text{maxSim}_2(D_4, D_5)] \\ &= [0.7 * 0.055 - (0.3) * 0] = 0.039 \end{aligned}$$

$$\begin{aligned} \text{MMR}(6) &= [0.7 * \text{Sim}_1(D_6, Q) - (1 - 0.7) * \text{maxSim}_2(D_6, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.084] = -0.025 \end{aligned}$$

$$\begin{aligned} \text{MMR}(8) &= [0.7 * \text{Sim}_1(D_8, Q) - (1 - 0.7) * \text{maxSim}_2(D_8, D_5)] \\ &= [0.7 * 0 - (0.3) * 0.008] = -0.003 \end{aligned}$$

Dari hasil perhitungan iterasi ke-6 bobot  $\text{ArgmaxMMR} = 0.039$  pada D4 atau kalimat ke-4, kemudian dipilih sebagai kalimat yang akan dijadikan kalimat ringkasan. Karena batas nilai maksimum antara 0 atau 1 maka iterasi akan berhenti ketika nilai maksimum 0.

Dari perhitungan MMR diatas maka di dapatkan hasil MMR tiap dokumen. Dapat dilihat pada tabel 4.8.

Tabel 4.8 Hasil iterasi MMR

iterasi ke	D1	D2	D3	D4	D5	D6	D7	D8	D9
1	<b>0.255</b>	0.096	0	0.039	0.096	0	0.099	0	0.050
2	-	0.072	0	0.029	0.096	0	<b>0.099</b>	0	0.025
3	-	0.072	0	0.029	<b>0.096</b>	0	-	0	0.025

4	-	<b>0.073</b>	-0.003	0.039	-	-0.025	-	-0.003	0.040
5	-	-	-0.003	0.039	-	-0.025	-	-0.003	<b>0.040</b>
6	-	-	-0.003	<b>0.039</b>	-	-0.025	-	-0.003	-

Berdasarkan perhitungan iterasi MMR maka dapat dilihat kalimat mana saja yang menjadi ringkasan. Kalimat yang akan dijadikan ringkasan adalah kalimat yang memiliki bobot tertinggi. Berikut adalah hasil bobot MMR maksimum.

Tabel 4.9 hasil bobot MMR maksimum iterasi MMR

Iterasi ke	Kalimat	Bobot ArgMax MMR
MMRMAX1	D1	0.255
MMRMAX2	D7	0.099
MMRMAX3	D5	0.096
MMRMAX4	D2	0.073
MMRMAX5	D9	0.040
MMRMAX6	D4	0.039

Dapat dilihat dari tabel diatas hasil bobot MMR maksimum yaitu pada D1, D7, D5, D2, D9, D4. Jadi yang menjadi kalimat peringkasan yaitu pada kalimat 1,7,5,9 dan 4.

#### 4.3.3.6 Analisa Query Expansion

*Query expansion* merupakan suatu teknik dengan menambahkan *keyword* baru kedalam *query* awal sehingga meningkatkan performasi pencarian. Tahap yang dilakukan yaitu :

1. lakukan perhitungan pembobotan *tf-idf*
2. ranking nilai bobot kata tersebut

3. lalu bobot kata yang tertinggi akan diambil sebagai *query expansion*

Adapun *Query expansion* pada dokumen yaitu : Dukung, Lengkap, Liter, Transmisi, Marg. Adapun nilai bobotnya

Dukung = 1.908

Lengkap = 1.908

Liter = 1.908

Marg = 1.908

Transmisi = 1.908

HASIL RINGKASAN :

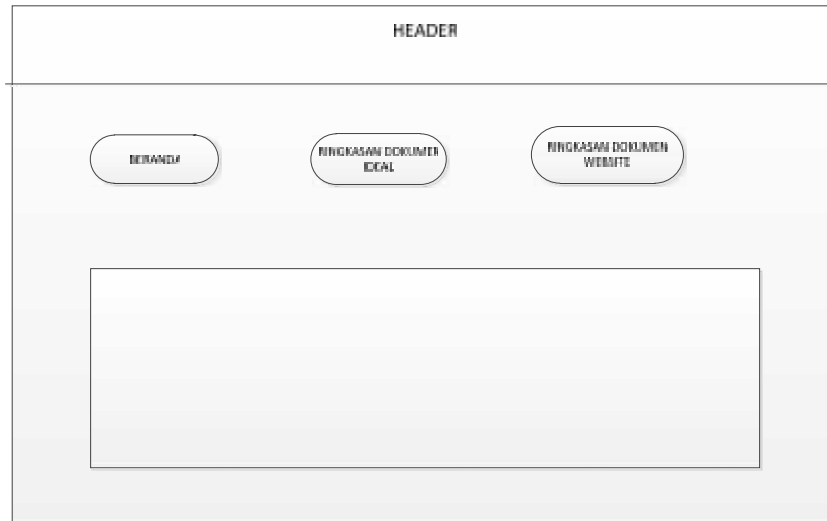
**“Tidak mau kalah, kepolisian amerika Gunakan Lamborghini Gallardo dukung lengkap transmisi liter marg”**

Tidak mau kalah dengan para polisi di daerah timur tengah, kini kepolisian Amerika Serikat (AS) khususnya Los Angeles dapat membusungkan dada mereka. Kehadiran mobil super LAPD yang diberikan seorang donatur guna mendukung sebuah kegiatan amal yang bermaksud untuk mendukung sebuah satuan udara milik LAPD. Lamborghini Gallardo adalah mobil super dengan menggunakan mesin V10 dengan kapasitas 5,0 liter atau mesin V10 5,2 liter. Pasalnya kini Los Angeles Police Department yang biasa dikenal dengan LAPD menambah skuad armada tempurnya dengan menggunakan supercar asal Italia, Lamborghini Gallardo. Terlihat mobil super ini memiliki desain khas lapd dengan dominan putih dan sudah dilengkapi dengan sirene serta kelengkapan lain layaknya mobil polisi. Beruntungnya adalah, LAPD tidak membeli kendaraan tersebut, melainkan hibah dari seorang pemilik perusahaan fiber optic di Amerika, Light Source bernama Nathale Marg dan Travis Marg.

Gambar 4.15 Hasil Ringkasan

## 4.4 Perancangan *Interface*

*Interface* sistem adalah sarana pengembangan sistem yang digunakan untuk membuat komunikasi yang lebih mudah, dan konsisten antara sistem dengan pemakainya. Penekanan *interface* meliputi tampilan yang baik, mudah dipahami dan tombol-tombol yang *familiar*. Berikut ini rancangan *interface* yang akan dibangun :



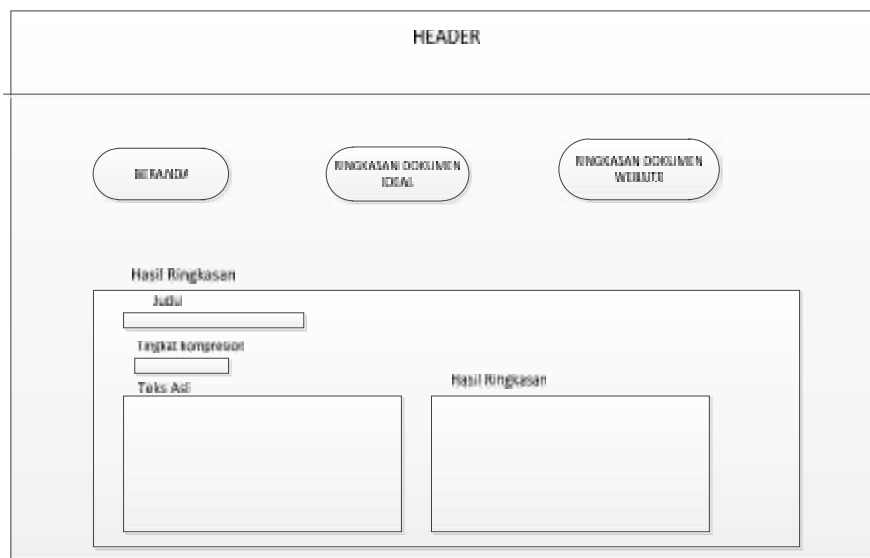
Gambar 4.16 Rancang *Interface* Sistem Peringkat

### 4.4.1 Rancangan *Interface* Menu Ringkasan Dokumen *Ideal*

Menu ringkasan dokumen *ideal* ini adalah halaman yang berfungsi untuk meringkas dokumen secara manual. Yang mana dokumen tersebut *dicopy* ke form "masukan teks" . kemudian isi judul dokumen tersebut. Atur tingkat *compression* Lalu setelah itu pilih tombol "proses". Berikut ini adalah rancangan *interface* menu ringkasan dokumen *ideal* :



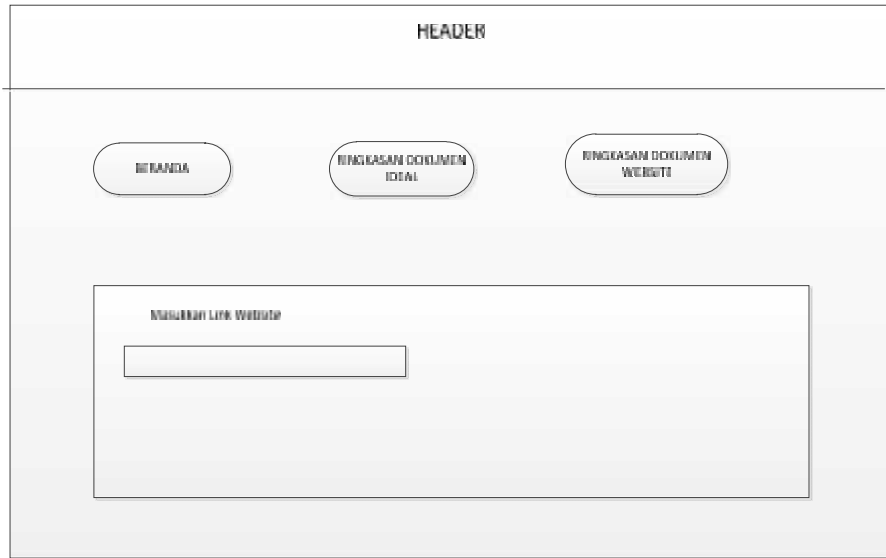
Gambar 4.17 Rancangan *Interface* Menu Ringkasan Dokumen *Ideal*



Gambar 4.18 Rancangan *Interface* Hasil Ringkasan Dokumen *Ideal*

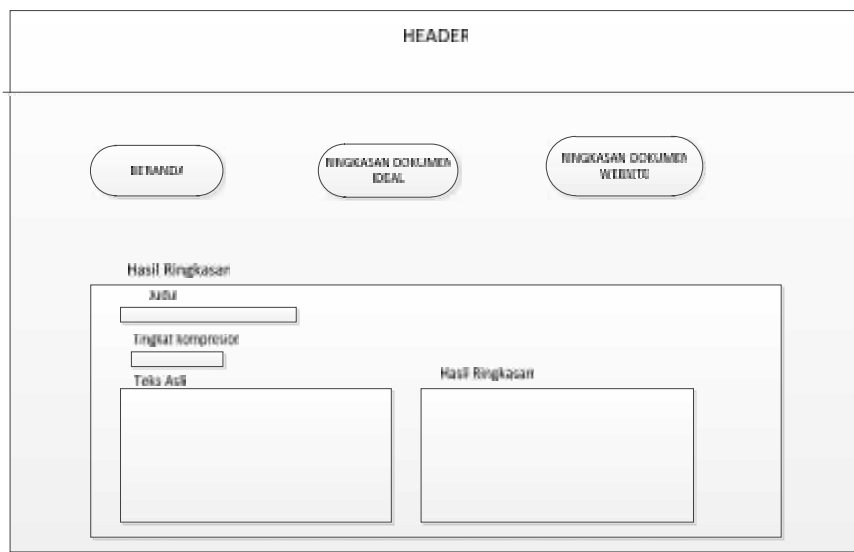
#### 4.4.2 Rancangan *Interface* Menu Ringkasan Dokumen *Website*

Menu ringkasan Dokumen *website* adalah halaman yang berfungsi untuk meringkas dokumen *online*. Masukkan judul dokumen (*query*) pada kolom ringkasan dokumen. Tombol *Compression* berfungsi untuk mengukur ringkasan pada dokumen tersebut. Tombol “*Search*” berfungsi untuk meringkas dokumen.



Gambar 4.19 Rancangan *Interface* Menu Ringkasan Dokumen *Website*

Tampilan yang akan muncul setelah menginputkan link *website*.



Gambar 4.20 Rancangan *Interface* hasil ringkasan Dokumen *Website*