BAB 3
ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Masalah

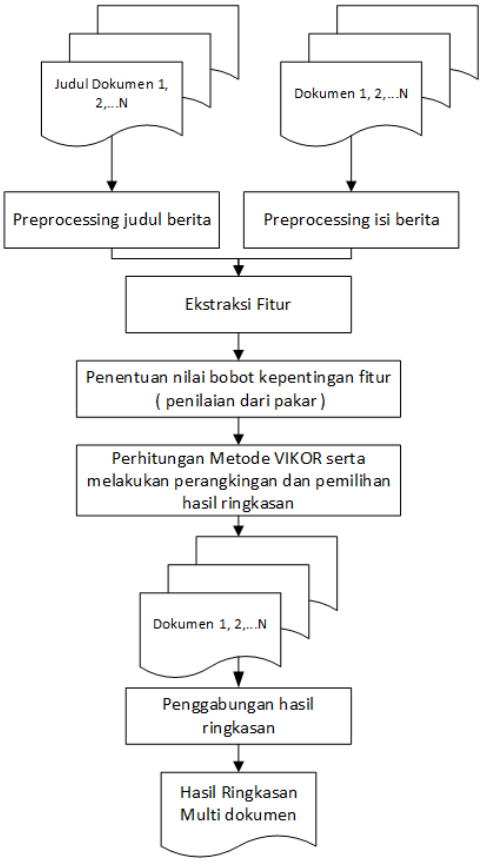
Berdasarkan uraian pada latar belakang masalah yang sudah dijelaskan sebelumnya pada penelitian ini terdapat sebuah masalah dalam peringkasan teks berita berbahasa indonesia yang memiliki topik yang sama dari berbagai sumber yang berbeda. Analisis masalah merupakan gambaran masalah yang akan diangkat dalam penelitian serta menjelaskan proses identifikasi masalah dan evaluasi mengenai sistem peringkasan otomatis yang digunakan pada peringkasan teks berita berbahasa indonesia. Maka dibutuhkan suatu metode untuk mengoptimalkan peringkasan teks otomatis agar menghasilkan ringkasan yang baik dan tidak menghilangkan isi inti dari berita tersebut.

Dalam jurnal penelitian sebelumnya yang membandingkan beberapa metode MCDM seperti *promethee, topsis, electre*, dan VIKOR menyatakan bahwa metode VIKOR adalah yang paling cocok untuk tugas pengambilan keputusan yang melibatkan pemilihan dari solusi retrofit untuk struktur tertentu, karena kemampuannya untuk menangani setiap jenis kriteria penilaian dan variabel, kejelasan hasil dan mengurangi efek dari sudut pandang subjektif pembuat keputusan [17]. Untuk penelitian multi dokumen,sebelumnya telah digunakan metode *Textrank*. Menghasilkan hasil peringkasan yang cukup baik. Hal ini ditunjukkan dari nilai akurasi yang di dapatkan sebesar 74%. Namun, metode *Texrank* memiliki kelemahan dalam penentuan tingkat kepentingan dari sebuah kalimat. Karena pada metode *Texrank* tidak menggunakan fitur ekstraksi yang digunakan untuk perhitungan kepentingan suatu kalimat berdasarkan fitur-fitur yang ada. Pada penelitian multi dokumen lainnya yang menggunakan salah satu metode MCDM yaitu TOPSIS yang juga merupakan metode perangkingan sama seperti VIKOR menghasilkan nilai yang dirasa kurang bagus yaitu *recall*, *precission,* dan *f-measure* sebesar 0.51, 0.52, 0.52 [6]. Hal ini disebabkan karena penentuan perangkingan tidak mempertimbangkan jarak relatif antar solusi idealnya. Karena nilai tertinggi dari alternatif TOPSIS tidak selalu menjadi solusi

ideal sehingga mempengaruhi hasil dari perangkingannya. Berdasarkan uraian diatas, maka akan dibuat sebuah penelitian yang bertujuan untuk mengimplementasikan metode VIKOR dalam peringkasan teks multi dokumen berbahasa indonesia.

3.2 Analisis Sistem

Sistem peringkasan teks otomatis yang akan dibangun pada penelitian ini digambarkan pada Gambar 3.4. Tahapan sistem peringkasan dimulai dari penginputan beberapa dokumen masukan, *prepocessing* judul yang kemudian akan dipakai pada saat perhitungan fitur kemiripan dengan judul,*preprocessing* isi, pembobotan fitur, perhitungan fitur-fitur ekstraksi. Hasil dari proses tersebut akan dilakukan proses klasifikasi dan perangkingan menggunakan metode VIKOR sehingga menghasilkan beberapa hasil ringkasan untuk kemudian digabungkan dengan TF-IDF dan *cosine similarity* sehingga menghasilkan ringkasan multi dokumen.



Gambar 3.1 Analisis Sistem

3.2.1 Analisis Data Masukan

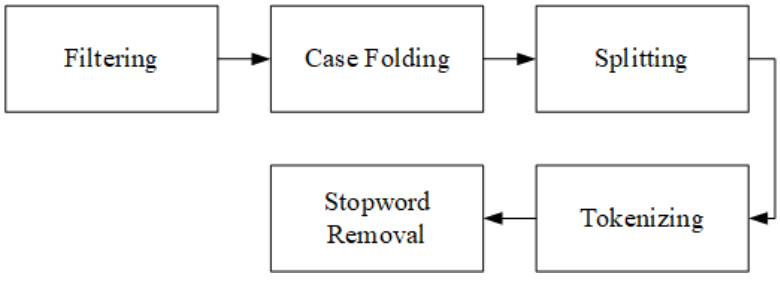
Data masukan yang diperlukan dalam menjalankan sistem ini sebelumnya telah dimasukkan kedalam database. Sehingga *user* bebas memilih berita mana yang akan di ringkas. Dalam penelitian ini dokumen yang dipakai adalah dokumen berita. Salah satu contoh data masukkan dapat dilihat pada Tabel 3.1.

Tabel 3.1 Data masukan

Olimpiade Pyeongchang 2018 Diserang Norovirus, Pemerintah Korsel Kerahkan Militer
PYEONGCHANG - Pemerintah Korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di Olimpiade Musim Dingin Pyeongchang. Pasalnya, para pekerja ini terkena norovirus. Virus menyebabkan seseorang seseorang mual dan muntah serta diare. Dan sejauh ini 32 pekerja telah dirawat karena norovirus. Untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina. "Untuk mengatasi kekurangan tenaga kerja keamanan yang diisolasi, sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil," bunyi pernyataan panpel Olimpiade Pyeongchang, Rabu 7 Februari 2018 dilansir Insidethegames. Mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih. "Untuk mencegah infeksi norovirus lebih lanjut, Kementerian Keamanan Pangan dan Obat-obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade. Pusat Pengendalian Penyakit Korsel juga memamtau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus." Dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet.

3.2.2 Analisis *Preprocessing*

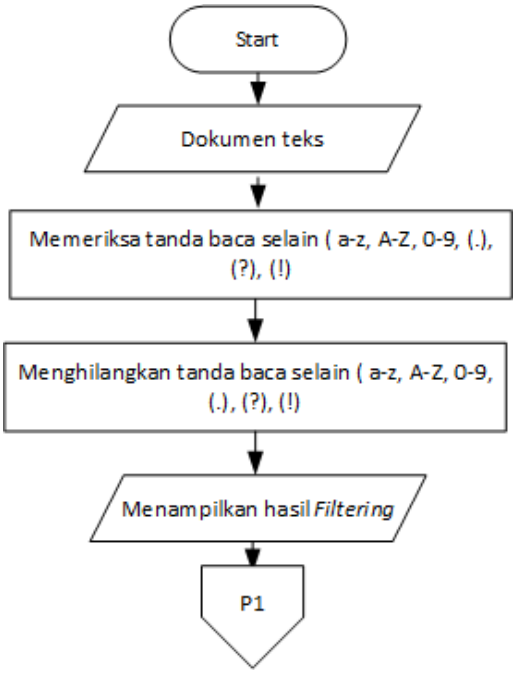
Tahap *preprocessing* merupakan tahapan untuk merapikan dan mempersiapkan dokumen yang akan digunakan dalam proses perhitungan pada ekstraksi fitur. *Preprocessing* pada penelitian ini terdiri dari berbagai tahapan yang digambarkan pada Gambar 3.5.



Gambar 3.2 *Preprocessing*

3.2.2.1 *Filtering*

Filtering adalah tahapan membuang tanda baca pada teks sehingga hanya terdapat kata a – z, A – Z, 0 – 9, spasi(“ ”) , (?), (!)dan tanda baca titik (“.”). Proses *filtering* ini bertujuan untuk menghilangkan karakter-karakter yang dianggap tidak berguna dan tidak diperlukan dalam tahapan selanjutnya. Proses *filtering* dapat dilihat pada gambar 3.3 :



Gambar 3.3 Proses P1 *Filtering*

Hasil *filtering* dapat dilihat pada Tabel 3.2.

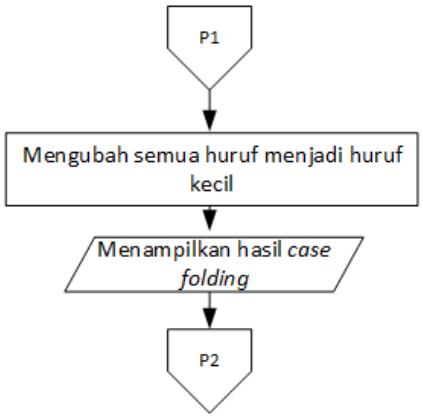
Tabel 3.2 Hasil *filtering*

Sebelum	Sesudah
PYEONGCHANG - Pemerintah Korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di Olimpiade Musim Dingin Pyeongchang. Pasalnya , para pekerja ini terkena norovirus. Virus menyebabkan seseorang seseorang mual dan muntah serta diare. Dan sejauh ini 32 pekerja telah dirawat karena norovirus. Untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina. "Untuk	PYEONGCHANG Pemerintah Korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di Olimpiade Musim Dingin Pyeongchang. Pasalnya para pekerja ini terkena norovirus. Virus menyebabkan seseorang seseorang mual dan muntah serta diare. Dan sejauh ini 32 pekerja telah dirawat karena norovirus. Untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina. Untuk

mengatasi kekurangan tenaga kerja keamanan yang diisolasi, sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil ," bunyi pernyataan panpel Olimpiade Pyeongchang, Rabu 7 Februari 2018 dilansir Insidethegames. Mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih. "Untuk mencegah infeksi norovirus lebih lanjut, Kementerian Keamanan Pangan dan Obat-obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade. Pusat Pengendalian Penyakit Korea juga memamtau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus." Dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet.	mengatasi kekurangan tenaga kerja keamanan yang diisolasi sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil bunyi pernyataan panpel Olimpiade Pyeongchang Rabu 7 Februari 2018 dilansir Insidethegames. Mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih. Untuk mencegah infeksi norovirus lebih lanjut Kementerian Keamanan Pangan dan Obat-obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade. Pusat Pengendalian Penyakit Korea juga memamtau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus. Dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet.

3.2.2.2 *Case Folding*

Case Folding adalah proses mengubah semua huruf pada teks dokumen menjadi karakter dengan huruf kecil atau huruf besar. Proses *case folding* dapat dilihat pada gambar 3.4 :



Gambar 3.4 Proses P2 *Case Folding*

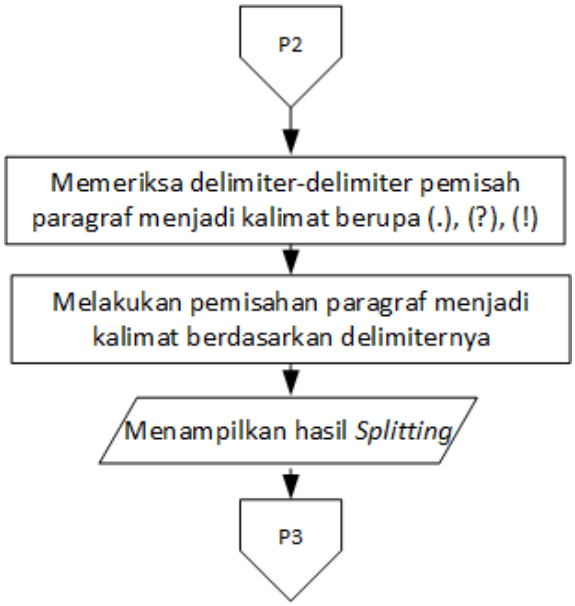
Contoh kata “PYEONGCHANG” diproses dengan *case folding* sehingga menjadi “pyeongchang”. Hasil *case folding* dapat dilihat pada Tabel 3.3.

Tabel 3.3 Hasil *Case Folding*

Sebelum	Sesudah
PYEONGCHANG Pemerintah Korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di Olimpiade Musim Dingin Pyeongchang. Pasalnya para pekerja ini terkena norovirus. Virus menyebabkan seseorang seseorang mual dan muntah serta diare. Dan sejauh ini 32 pekerja telah dirawat karena norovirus. Untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina. Untuk mengatasi kekurangan tenaga kerja keamanan yang diisolasi, sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil, bunyi pernyataan panpel Olimpiade Pyeongchang, Rabu 7 Februari 2018 dilansir Insidethegames. Mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih. Untuk mencegah infeksi norovirus lebih lanjut Kementerian Keamanan Pangan dan Obat obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade. Pusat Pengendalian Penyakit Korea juga memamtau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus. Dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet.	pyeongchang pemerintah korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di olimpiade musim dingin pyeongchang. pasalnya para pekerja ini terkena norovirus. virus menyebabkan seseorang seseorang mual dan muntah serta diare. dan sejauh ini 32 pekerja telah dirawat karena norovirus. untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina. untuk mengatasi kekurangan tenaga kerja keamanan yang diisolasi sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil bunyi pernyataan panpel olimpiade pyeongchang rabu 7 februari 2018 dilansir insidethegames. mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih. untuk mencegah infeksi norovirus lebih lanjut kementerian keamanan pangan dan obat obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade. pusat pengendalian penyakit korea juga memantau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus. dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet.

3.2.2.3 *Splitting*

Splitting adalah pemisahan teks dalam bentuk satuan kalimat berdasarkan delimiter-delimiter tertentu. Delimiter tersebut adalah tanda titik (“.”), tanda seru (“!”) , dan tanda tanya (“?”). proses *splitting* dapat dilihat pada gambar 3.5 :



Gambar 3.5 Proses P3 *Splitting*

Hasil dari *splitting* kalimat dapat dilihat pada Tabel 3.4.

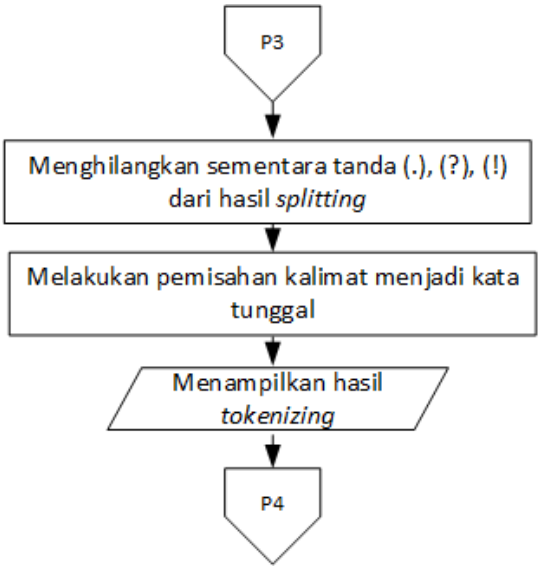
Tabel 3.4 Hasil *Splitting*

No	*Splitting* kalimat
1	pyeongchang pemerintah korsel mengerahkan militer untuk mengisi kekosongan pekerja yang akan bertugas di olimpiade musim dingin pyeongchang
2	pasalnya para pekerja ini terkena norovirus
3	virus menyebabkan seseorang mual dan muntah serta diare
4	dan sejauh ini 32 pekerja telah dirawat karena norovirus
5	untuk mencegah menyebarnya virus ini sebanyak 1200 pekerja telah dikarantina
6	untuk mengatasi kekurangan tenaga kerja keamanan yang diisolasi sekitar 900 personel militer telah ditugaskan untuk mengambil alih pekerjaan keamanan sipil bunyi pernyataan panpel olimpiade pyeongchang rabu 7 Februari 2018 dilansir insidethegames
7	mereka ini nantinya akan bekerja di 20 tempat pertandigan sampai semua pekerja dinyatakan pulih

No	*Splitting* kalimat
8	untuk mencegah infeksi norovirus lebih lanjut kementerian keamanan pangan dan obat obatan akan menginspeksi kebersihan di restoran dan semua fasilitas makanan yang terkait dengan olimpiade
9	pusat pengendalian penyakit korea juga memamtau pasien diare dan meminta semua orang untuk mengikuti langkah pencegahan norovirus
10	dalam pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet

3.2.2.4 *Tokenizing*

Tokenizing adalah proses pemisahan deretan kata dari suatu kalimat atau paragraf menjadi kata tunggal. Proses *tokenizing* dapat dilihat pada gambar 3.6 :



Gambar 3.6 Proses P4 *Tokenizing*

Hasil *tokenizing* dapat dilihat pada Tabel 3.5.

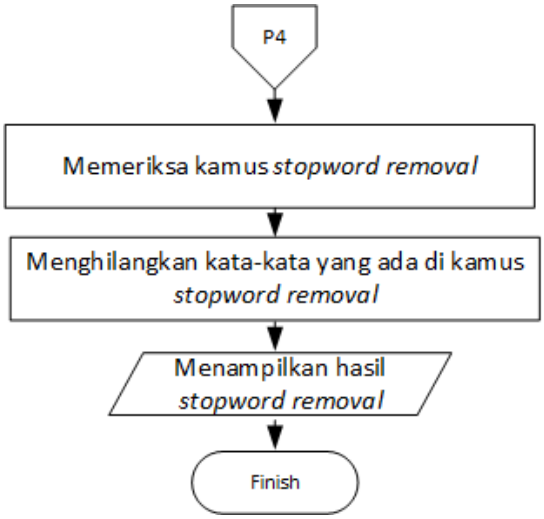
Tabel 3.5 Hasil *Tokenizing*

No	*Tokenizing* kata				
1	pyeongchang	pemerintah	korsel	mengerahkan	militer
	untuk	mengisi	kekosongan	pekerja	yang
	akan	bertugas	di	olimpiade	musim
	dingin	pyeongchang			
2	pasalnya	para	pekerja	ini	terkena
	norovirus				
3	virus	menyebabkan	seseorang	mual	dan

No	*Tokenizing* kata				
	muntah	serta	diare		
4	dan	sejauh	ini	32	pekerja
	telah	dirawat	karena	norovirus	
5	untuk	mencegah	menyebarnya	virus	ini
	sebanyak	1200	pekerja	telah	dikarantina
6	untuk	mengatasi	kekurangan	tenaga	kerja
	keamanan	yang	diisolasi	sekitar	900
	personel	militer	telah	ditugaskan	untuk
	mengambil	alih	pekerjaan	keamanan	sipil
	bunyi	pernyataan	panpel	olimpiade	pyeongchang
	rabu	7	februari	2018	dilansir
	insidethegames				
7	mereka	ini	nantinya	akan	bekerja
	di	20	tempat	pertandingan	sampai
	semua	pekerja	dinyatakan	pulih	
8	untuk	mencegah	infeksi	norovirus	lebih
	lanjut	kementerian	keamanan	pangan	dan
	obat	obatan	akan	menginspeksi	kebersihan
	di	restoran	dan	semua	fasilitas
	makanan	yang	terkait	dengan	olimpiade
9	pusat	pengendalian	penyakit	korea	juga
	memamtau	pasien	diare	dan	meminta
	semua	orang	untuk	mengikuti	langkah
	pencegahan	norovirus			
10	dalam	pernyataan	itu	panpel	juga
	memberikan	saran	terhadap	masyarakat	dntuk
	mencuci	tangan	selama	30	detik
	dengan	air	yang	mengalir	dan
	sabun	setelah	pergi	ke	toilet

3.2.2.5 *Stopword Removal*

Stopword removal adalah proses untuk menghilangkan kata – kata yang tidak memiliki makna atau kata yang kurang memiliki pengaruh dan sering muncul. Kamus kata diambil dari penelitian Fadillah Z Tala yang digunakan untuk proses *stopword removal* digunakan untuk melihat kata mana saja yang akan dihilangkan[6]. Kata-kata didalam kamus tersebut kemudian diseleksi kembali dengan mempertimbangkan makna kata tersebut dalam sebuah dokumen berita. Seperti kata "sebanyak" dan "sebesar" yang dirasa mempunyai makna dalam menyatakan jumlah. Proses *stopword removal* dapat dilihat pada gambar 3.7 dan Tabel 3.6 :



Gambar 3.7 Proses *Stopword Removal*

Hasil dari *stopword removal* dapat dilihat pada tabel 3.6

Tabel 3.6 Proses *Stopword Removal*

No	*Tokenizing* kata				
1	pyeongchang	pemerintah	korsel	mengerahkan	militer
	untuk	mengisi	kekosongan	pekerja	yang
	akan	bertugas	di	olimpiade	musim
	dingin	pyeongchang			
2	pasalnya	para	pekerja	ini	terkena
	norovirus				
3	virus	menyebabkan	seseorang	mual	dan
	muntah	serta	diare		

No	*Tokenizing* kata				
4	dan	sejauh	ini	32	pekerja
	telah	dirawat	karena	norovirus	
5	untuk	mencegah	menyebarnya	virus	ini
	sebanyak	1200	pekerja	telah	dikarantina
6	untuk	mengatasi	kekurangan	tenaga	kerja
	keamanan	yang	diisolasi	sekitar	900
	personel	militer	telah	ditugaskan	untuk
	mengambil	alih	pekerjaan	keamanan	sipil
	bunyi	pernyataan	panpel	olimpiade	pyeongchang
	rabu	7	februari	2018	dilansir
	insidethegames				
7	mereka	ini	nantinya	akan	bekerja
	di	20	tempat	pertandingan	sampai
	semua	pekerja	dinyatakan	pulih	
8	untuk	mencegah	infeksi	norovirus	lebih
	lanjut	kementerian	keamanan	pangan	dan
	obat	obatan	akan	menginspeksi	kebersihan
	di	restoran	dan	semua	fasilitas
	makanan	yang	terkait	dengan	olimpiade
9	pusat	pengendalian	penyakit	korea	juga
	memamtau	pasien	diare	dan	meminta
	semua	orang	untuk	mengikuti	langkah
	pencegahan	norovirus			
10	dalam	pernyataan	itu	panpel	juga
	memberikan	saran	terhadap	masyarakat	untuk
	mencuci	tangan	selama	30	detik
	dengan	air	yang	mengalir	dan
	sabun	setelah	pergi	ke	toilet

Hasil dari proses *stopword removal* dapat dilihat pada Tabel 3.7.

Tabel 3.7 Hasil Proses *Stopword Removal*

No	*Splitting* kalimat
1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang

No	*Splitting* kalimat
2	pasalnya pekerja terkena norovirus
3	virus menyebabkan seseorang mual muntah diare
4	sejauh 32 pekerja dirawat karena norovirus
5	mencegah menyebarnya virus sebanyak 1200 pekerja dikarantina
6	mengatasi kekurangan tenaga kerja keamanan diisolasi sekitar 900 personel militer ditugaskan mengambil alih pekerjaan keamanan sipil bunyi pernyataan panpel olimpiade pyeongchang rabu 7 Februari 2018 dilansir insidethegames
7	mereka nantinya bekerja 20 tempat pertandigan pekerja dinyatakan pulih
8	mencegah infeksi norovirus lanjut kementerian keamanan pangan obat obatan menginspeksi kebersihan restoran fasilitas makanan terkait olimpiade
9	pusat pengendalian penyakit korsel memamtau pasien diare meminta orang mengikuti langkah pencegahan norovirus
10	pernyataan itu panpel juga memberikan saran terhadap masyarakat untuk mencuci tangan selama 30 detik dengan air yang mengalir dan sabun setelah pergi ke toilet

3.2.3 Analisis Ekstraksi Fitur

Ekstraksi fitur adalah proses pengolahan kalimat menggunakan fitur-fitur sehingga menghasilkan nilai-nilai fitur yang akan digunakan untuk peringkasan dokumen. Proses ini dilakukan setelah melalui proses *preprocessing*. Fitur-fitur yang digunakan adalah posisi kalimat, kemiripan dengan judul, koneksi antar kalimat, kata numerik.

3.2.3.1 Posisi Kalimat (F1)

Fitur ini dipakai berdasarkan asumsi bahwa kalimat yang paling penting pada sebuah dokumen adalah kalimat pertama dan kalimat terakhir. Hasil ini didapatkan berdasarkan rumus (2.1). contoh perhitungan fitur posisi kalimat dapat dilihat pada tabel 3.8.

$$Posisi\ kalimat\ = Max\left(\frac{1}{i}, \frac{1}{n-i+1}\right)$$

Ket : i = urutan kalimat pada dokumen

n = jumlah keseluruhan kalimat pada dokumen

Tabel 3.8 Hasil Perhitungan Fitur Posisi Kalimat

Kalimat	Posisi kalimat
S1	$Posisi\ kalimat\ = Max\left(\frac{1}{1}, \frac{1}{10-1+1}\right)$ = 1
S2	$Posisi\ kalimat\ = Max\left(\frac{1}{2}, \frac{1}{10-2+1}\right)$ = 0,5
S3	$Posisi\ kalimat\ = Max\left(\frac{1}{3}, \frac{1}{10-3+1}\right)$ = 0,3333
S4	$Posisi\ kalimat\ = Max\left(\frac{1}{4}, \frac{1}{10-4+1}\right)$ = 0,25
S5	$Posisi\ kalimat\ = Max\left(\frac{1}{5}, \frac{1}{10-5+1}\right)$ = 0,2
S6	$Posisi\ kalimat\ = Max\left(\frac{1}{6}, \frac{1}{10-6+1}\right)$ = 0,2
S7	$Posisi\ kalimat\ = Max\left(\frac{1}{7}, \frac{1}{10-7+1}\right)$ = 0,25
S8	$Posisi\ kalimat\ = Max\left(\frac{1}{8}, \frac{1}{10-8+1}\right)$ = 0,3333
S9	$Posisi\ kalimat\ = Max\left(\frac{1}{9}, \frac{1}{10-9+1}\right)$ = 0,5
S10	$Posisi\ kalimat\ = Max\left(\frac{1}{10}, \frac{1}{10-10+1}\right)$ = 1

3.2.3.2 Kata Kunci (F2)

Awalnya, kata kunci didefinisikan sebagai kata non-isyarat dengan bobot terukur yang tertinggi dan bobot kunci yang digunakan menjadi frekuensi kemunculan di sebuah dokumen. Dalam penelitian ini, fitur kata kunci yang diekstrak didapatkan dari ≤ 10 kata yang sering muncul dalam sebuah dokumen. Setelah kata kunci didapatkan, setiap token dalam kalimat yang sama dengan kata kunci akan diberi nilai 0,1 dan yang tidak sama akan diberi nilai 0.

Daftar kata kunci dapat dilihat pada Tabel 3.9 dan pembobotan kata kunci dapat dilihat pada Tabel 3.10 :

Tabel 3.9 Kata Kunci

No	Kata	Jumlah
1	Pekerja	5
2	Norovirus	4
3	Pyeongchang	3
4	Olimpiade	3
5	Keamanan	3
6	Virus	2
7	Diare	2

No	Kata	Jumlah
8	Pernyataan	2
9	Militer	2
10	Mencegah	2

Setelah didapatkan ≤10 kata yang sering muncul dalam dokumen, kemudian dilakukan perhitungan nilai bobot pada kata kunci pada setiap kalimatnya. Jika didalam kalimat terdapat kata kunci tersebut maka bobot setiap kata kunci yang ada di kalimat tersebut akan ditambahkan 0,1 setiap kemunculan kata kunci.

Tabel 3.10 Pembobotan Kata Kunci

Kalimat	$\sum total\ bobot$
S1	0,5
S2	0,2
S3	0,2
S4	0,2
S5	0,3
S6	0,6
S7	0,1
S8	0,4
S9	0,2
S10	0,1

3.2.3.3 Kemiripan Kalimat dengan Judul (F3)

Judul dalam sebuah berita merupakan pemadatan dari sebuah berita dan menggambarkan isi dari berita tersebut. Pada penelitian ini penggunaan dokumen yang digunakan adalah berita yang berjudul "Olimpiade Pyeongchang 2018 Diserang Norovirus, pemerintah korsel kerahkan militer". Judul tersebut akan melewati proses *preprocessing* untuk selanjutnya dibandingkan dengan kalimat-kalimat hasil *preprocessing*. Hasil *preprocessing* dapat dilihat pada Tabel 3.11.

Tabel 3.11 *Preprocessing* Judul

No	Token
1	olimpiade
2	pyeongchang
3	2018

No	Token
4	diserang
5	norovirus
6	pemerintah
7	korsel
8	kerahkan
9	militer

Setelah judul melalui tahapan *preprocessing*, kemudian dilakukan perhitungan kecocokan dari setiap kata pada judul tersebut dihitung dengan rumus 2.2. Contoh perhitungan fitur kata kunci untuk kalimat S1 dapat dilihat pada Tabel 3.12 :

$$Scoref3\ (s) = \frac{|kata\ di\ judul \cap kata\ di\ kalimat|}{|kata\ di\ judul \cup kata\ di\ kalimat|}$$

Tabel 3.12 Token Kalimat 1

Kalimat	Kata
1	pyeongchang
1	pemerintah
1	korsel
1	mengerahkan
1	militer
1	mengisi
1	kekosongan
1	pekerja
1	bertugas
1	olimpiade
1	musim
1	dingin
1	pyeongchang

Pada kalimat S1 didapat 5 kesamaan kata dengan judul berita yaitu, “olimpiade”,”pyeongchang”,”pemerintah”,”korsel”, dan ”militer”. Maka perhitungannya :

$$score(S1) = \frac{5}{22} = 0{,}2273$$

Dimana : 5 = kata dikalimat yang sama dengan judul

22 = jumlah gabungan kata antar kalimat

Untuk hasil perhitungan nilai kemiripan dengan judul untuk semua kalimat dapat dilihat pada Tabel 3.13 :

Tabel 3.13 Bobot Kemiripan Semua Kalimat dengan Judul

Kalimat	Bobot kemiripan
S1	$score(S1) = \frac{5}{22} = 0{,}2273$
S2	$score(S2) = \frac{1}{13} = 0{,}0796$
S3	$score(S3) = \frac{0}{15} = 0$
S4	$score(S4) = \frac{1}{13} = 0{,}0455$
S5	$score(S5) = \frac{0}{17} = 0$
S6	$score(S6) = \frac{2}{36} = 0{,}1143$
S7	$score(S7) = \frac{0}{19} = 0$
S8	$score(S8) = \frac{2}{23} = 0{,}08696$
S9	$score(S9) = \frac{2}{21} = 0{,}09524$
S10	$score(S10) = \frac{0}{22} = 0$

3.2.3.4 Bobot Koneksi Antar Kalimat (F4)

Koneksi antar kalimat adalah menghitung banyak kata yang sama dari satu kalimat dengan kalimat yang lain dalam satu dokumen. Perhitungan koneksi antar kalimat menggunakan rumus 2.4.

$$scoref4\ (Si)\ = \sum koneksi\ antar\ kalimat$$

Hasil perhitungan bobot koneksi antar kalimat dapat dilihat pada tabel 3.14

Tabel 3.14 Bobot Koneksi Antar Kalimat

Kalimat	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	0	1	0	1	1	3	1	1	1	0
S2	1	0	0	2	1	0	1	1	1	0

Kalimat	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S3	0	0	0	0	1	0	0	0	1	0
S4	1	2	0	0	1	0	1	1	1	0
S5	1	1	1	1	0	0	1	1	0	0
S6	3	0	0	0	0	0	0	2	0	2
S7	1	1	0	1	1	0	0	0	0	0
S8	1	1	0	1	1	2	0	0	1	0
S9	1	1	1	1	0	0	0	1	0	0
S10	0	0	0	0	0	2	0	0	0	0
$\sum total$	9	7	2	7	6	7	4	7	5	2

3.2.3.5 Kalimat yang Mengandung Kata Numerik (F5)

Pada peringkasan teks khususnya teks berita haruslah mempertimbangkan kata yang mengandung unsur numerik. Karena dalam kalimat yang berisi numerik terdapat kalimat yang penting. Dengan menggunakan rumus 2.5.

$$Scoref5\ (Si) = \frac{banyak\ angka\ dalam\ kalimat(s)}{panjang(s)}$$

Perhitungan fitur kalimat yang mengandung kata numerik dapat dilihat pada tabel 3.15.

Tabel 3.15 Data Numerik Pada Kalimat

Kalimat	Bobot
S1	$Score\ (S1) = \frac{0}{13} = 0$
S2	$Score\ (S2) = \frac{0}{4} = 0$
S3	$Score\ (S3) = \frac{0}{5} = 0$
S4	$Score\ (S4) = \frac{1}{4} = 0.25$
S5	$Score\ (S5) = \frac{1}{6} = 0{,}1667$
S6	$Score\ (S6) = \frac{3}{26} = 0{,}1154$
S7	$Score\ (S7) = \frac{1}{5} = 0{,}2$
S8	$Score\ (S8) = \frac{0}{13} = 0$

Kalimat	Bobot
S9	$Score\ (S9) = \frac{0}{12} = 0$
S10	$Score\ (S10) = \frac{1}{13} = 0{,}0769$

3.2.3.6 Frekuensi Kata (F6)

Kemunculan kata dicek keberadaannya pada setiap dokumen, karena kebermaknaan suatu kata dapat ditentukan dengan seberapa banyak kata tersebut muncul dalam setiap dokumen. Untuk itu digunakan sebuah metode yang dapat mengekstrak data frekuensi kata dalam dokumen yaitu tf-isf. Tf-isf merupakan sebuah metode statistical yang digunakan untuk mengevaluasi kebermaknaan kata dalam sebuah kalimat pada dokumen. Metode ini diterapkan setelah kalimat melalui proses *stopword* sudah dihilangkan dari kalimat. Rumus tf-isf dapat dihitung dengan rumus 2.6 dan 2.7:

$$isf = log\left(\frac{N}{n_i}\right)$$

Keterangan : N = jumlah keseluruhan kalimat pada dokumen

ni = jumlah kata yang mengandung kata ke-i

$$frekuensi\ kata = Tf_i * isf$$

Keterangan : Tf_i = jumlah kemunculan kata dalam kalimat

Berikut contoh perhitungan isf yang merupakan nilai yang akan dibutuhkan pada perhitungan tf-isf :

Tabel 3.16 Contoh Perhitungan Isf pada Kalimat

Kata	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	ni	Log(N/ni)
pyeongchang	2	0	0	0	0	1	0	0	0	0	2	0.6989
pemerintah	1	0	0	0	0	0	0	0	0	0	1	1
korsel	1	0	0	0	0	0	0	0	1	0	1	1
mengerahkan	1	0	0	0	0	0	0	0	0	0	1	1
militer	1	0	0	0	0	1	0	0	0	0	2	0.6989

Kemunculan kata yang sama pada kalimat yang sama diabaikan dalam tahapan isf. Sebagai contoh kata “pyeongchang” muncul 2 kali dalam kalimat

pertama , namun nilai kemunculan kata tersebut hanya dihitung satu karena dalam tahap isf hanya mengecek kemunculan kata dalam dokumen maka nilai ni = 1. Karena kata "pyeongchang" juga muncul pada kalimat "S6" maka nilai ni = 2. Berikut adalah perhitungan nilai isf untuk kata "pyeongchang dengan menggunakan rumus 2.6 :

$$pyeongchang = log\left(\frac{10}{2}\right)$$

$$= 0.6989$$

Dimana : 10 : jumlah keseluruhan kalimat dalam dokumen

2 : jumlah kata yang mengandung kata ke-*i* dalam kalimat

Setelah didapatkan nilai isf dari setiap kata, selanjutnya dilakukan perhitungan dengan rumus tf*isf. Berikut cara perhitungan TF-ISF untuk kata "pyeongchang" dengan menggunakan rumus 2.7 :

$$tf = 3 * 0.6989$$

$$= 2.0967$$

Keterangan : Tf = jumlah kemunculan kata pada kalimat tertentu

Setelah mendapatkan nilai dari masing-masing kata, maka nilai dari frekuensi kata yang muncul pada masing-masing kalimat ditentukan dengan menambahan setiap nilai hasil dari perhitungan tf-isf. Berikut hasil penjumlahan dari nilai tf-isf pada setiap kalimat :

Tabel 3.17 Tabel Perhitungan TF-ISF pada Setiap Kalimat

Kalimat	TF-ISF
S1	10.6198
S2	2.6989
S3	4.3979
S4	2.6989
S5	4.6989
S6	23.7167
S7	4.3010
S8	12.3187
S9	10.7959
S10	12.3979

3.2.3.7 Sentralitas (F7)

Sentralitas dalam sebuah kalimat di dokumen merupakan kemiripan kata antara kalimat yang satu dengan yang lainnya. Bila sebuah kalimat memiliki sentralitas yang tinggi,maka kalimat tersebut memiliki kemungkinan tertinggi untuk dimasukan kedalam hasil ringkasan. Sentralitas digunakan untuk mencari hubungan antar kalimat-kalimat yang ada di dokumen sampai ditemukan nilai sentralitas tertinggi yang dimiliki oleh sebuah kalimat. Nilai sentralitas dihitung dengan menggunakan rumus 2.8. Hasil perhitungan nilai sentralitas dari setiap dokumen dengan dokumen lainnya dapat dilihat pada tabel 3.18 :

Tabel 3.18 Nilai Sentralitas Seluruh Kalimat

Kalimat	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	0	0.0588	0	0.0588	0.0526	0.0769	0.0555	0.0370	0.04	0
S2	0.0588	0	0	0.25	0.1	0	0.1111	0.0556	0.0625	0
S3	0	0	0	0	0.0909	0	0	0	0.0588	0
S4	0.0588	0.25	0	0	0.1	0	0.1111	0.0556	0.0625	0
S5	0.0526	0.1	0.0909	0.1	0	0	0.0909	0.05	0	0
S6	0.0769	0	0	0	0	0	0	0.05	0	0,0513
S7	0.0555	0.1111	0	0.1111	0.0909	0	0	0	0	0
S8	0.0370	0.0556	0	0.0556	0.05	0.05	0	0	0.0385	0
S9	0.04	0.0625	0.0588	0.0625	0	0,0513	0	0.0385	0	0
S10	0	0	0	0	0	0	0	0,0513	0	0

Pada kalimat S1 didapatkan kesamaan kata dengan kalimat S2 sebanyak 1 kata yaitu “pekerja” dan jumlah kata yang ada dikalimat S1 dan S2 adalah 17. Berikut contoh perhitungan nilai sentralitas antara kalimat S1 dengan S2 :

$$sentralitas = \frac{1}{17}$$

$$= 0.0588$$

Dimana : 1 : merupakan nilai kesamaan antar kalimat S1 dan S2

17 : jumlah dari seluruh kata pada kalimat S1 dan S2

Jika nilai sentralitas setiap kalimat di dalam dokumen sudah didapatkan. Selanjutnya akan diambil nilai sentralitas terbesar dari setiap kalimat. Tabel nilai sentralitas maks pada setiap kalimat dapat dilihat pada tabel 3.19:

Tabel 3.19 Nilai Sentralitas Maks pada Setiap Kalimat

Kalimat	Nilai sentralitas maks	Terhadap Kalimat-
S1	0.07692	S6
S2	0.25	S4
S3	0.0909	S5
S4	0.25	S2
S5	0.1	S2
S6	0.07692	S1
S7	0.1111	S2 dan S4
S8	0.0556	S2 dan S4
S9	0.0625	S2 dan S4
S10	0.05128	S8

Jika proses ekstraksi fitur sudah selesai dilakukan. Maka semua nilai dari fitur ekstraksi akan dipetakan kedalam matriks. Nilai tersebut akan digunakan pada salah satu tahapan dalam perhitungan metode VIKOR yaitu tahapan normalisasi matriks. Tabel nilai dari hasil perhitungan fitur ekstraksi dapat dilihat pada tabel 3.20.

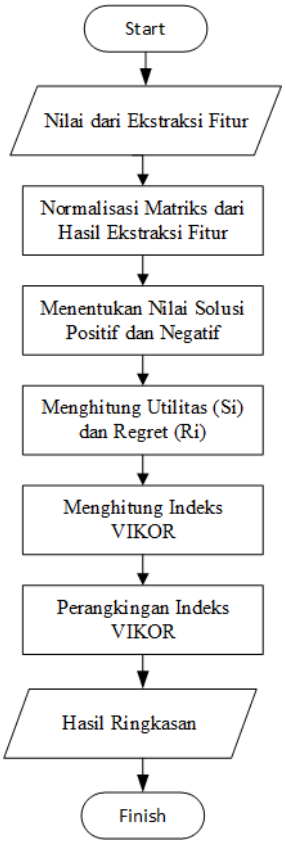
Tabel 3.20 Hasil Ekstraksi Seluruh Fitur

	F1	F2	F3	F4	F5	F6	F7
S1	1	0,5	0,2727	9	0	10.6198	0.07692
S2	0,5	0,2	0,07692	7	0	2.6989	0.25
S3	0,3333	0,2	0	2	0	4.3979	0.0909
S4	0,25	0,2	0,07692	7	0.25	2.6989	0.25
S5	0,2	0,3	0	6	0,1667	4.6989	0.1
S6	0,2	0,6	0,1143	7	0,1154	23.7167	0.07692
S7	0,25	0,1	0	4	0,2	4.3010	0.1111
S8	0,3333	0,4	0,0869	7	0	12.3187	0.0556
S9	0,5	0,2	0,0952	5	0	10.7959	0.0625
S10	1	0,1	0	2	0,0769	12.3979	0.05128

Hasil dari perhitungan bobot dari semua ektraksi fitur ini digunakan untuk kriteria yang nantinya akan dipakai dalam perhitungan metode VIKOR untuk menghasilkan sebuah ringkasan teks.

3.2.4 Perhitungan Metode VIKOR

VIKOR merupakan salah satu metode *multi criteria decision making* (MCDM) yang berfokus pada perangkingan dan pemilihan berdasarkan dari hasil alternatif dan kriteria yang bertentangan yang menghasilkan solusi yang mungkin paling dekat dengan solusi ideal. Berikut adalah tahapan-tahapan penerapan metode VIKOR :



Gambar 3.8 Alur Metode VIKOR

3.2.4.1 Bobot Kepentingan Fitur

Perhitungan bobot fitur dilakukan sebagai data masukkan untuk fungsi subjektif metode VIKOR. Untuk bobot fitur sendiri ditentukan oleh pakar [6], dimana pakar dalam bidang peringkasan teks otomatis berbahasa Indonesia ini berasal dari latar belakang pendidikan sastra bahasa Indonesia yaitu Mahmud Fasya M.A, pakar 2 Dina Novalia S.S, dan pakar ke 3 Tri Ratna Nigrum S,Pd. Tingkat kepentingan fitur atau bobot fitur yang digunakan dalam penelitian ini diberikan bobot sebagai berikut.

Tabel 3.21 Pembobotan Fitur Ekstraksi oleh Pakar

No	Fitur	Penjelasan	Bobot (%) oleh Mahmud Fasya M.A	Bobot (%) oleh Dina Novalia	Bobot (%) oleh Tri Ratna Nigrum
1	Posisi Kalimat	Dalam dokumen ilmiah, kalimat penting dalam suatu dokumen biasanya berada pada bagian awal dan akhir	14,29%	15%	20%
2	Kata Penting/Kata kunci	Tingginya tingkat kemunculan sebuah kata tertentu pada sebuah dokumen, memperlihatkan bahwa kalimat yang mengandung kata tersebut penting.	21,43%,	20%	20%
3	Kemiripan dengan Judul	Judul merupakan sekumpulan kata yang arah penting mengenai konsep dari sebuah dokumen, maka dari itu jika pada sebuah kalimat memiliki kemiripan dengan judulnya itu bisa jad menandakan kalimat itu penting.	25%,	15%	30%
4	Koneksi antar Kalimat	Memperhitungkan banyaknya kata yang sama dalam tiap kalimat dalam sebuah dokumen.	7,14%.	10%	5%
5	Kalimat yang mengandung kata numerik	Mempertimbangkan data numerik (kata yang mengandung angka), karena dalam sebuah kalimat yang memiliki data numerik	3,57%	5%	5%

No	Fitur	Penjelasan	Bobot (%) oleh Mahmud Fasya M.A	Bobot (%) oleh Dina Novalia	Bobot (%) oleh Tri Ratna Nigrum
		biasanya menandakan bahwa kalimat tersebut penting.			
6	Frekuensi Kata/ TF-ISF	Kebermaknaan sebuah kata dapat dinilai dari seberapa sering kata tersebut muncul pada sebuah kalimat	17,86%,	25%	10%
7	Sentralitas	Sentralitas dari sebuah kalimat menyatakan kemiripannya pada kalimat lain, jadi bila sebuah kalimat memiliki sentralitas yang lebih tinggi dari kalimat lainnya, maka kalimat tersebut merupakan kandidat terbaik.	10,71%	10%	10%

3.2.4.2 Normalisasi Matriks

Matriks keputusan yang telah didapatkan dari beberapa kriteria yang diambil dari perhitungan bobot ekstraksi fitur akan di normalisasi untuk menghasilkan matriks yang baru dari bobot fitur-fitur sebelumnya. Perhitungan normalisasi matriks menggunakan rumus 2.9.

$$f_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m}\left(x_{ij}^{2}\right)}}, i = 1, \dots, n; j = 1, \dots, m$$

Keterangan : x_{ij} = elemen matriks dari kolom *i* baris ke *j*

n = jumlah kalimat

m = jumlah fitur

Berikut contoh dari perhitungan normalisasi matriks pada elemen matriks x_{11} :

$$f_{11} = \frac{1}{\sqrt{((1^2) + (0.5^2) + (0.2727^2) + (0.9^2) + (0^2) + (10.6198^2) + (0.0769^2))}}$$

$$f_{11} = \frac{1}{\sqrt{(1 + 0.25 + 0.07436 + 81 + 0 + 112.7801 + 0.00592)}}$$

$$f_{11} = \frac{1}{\sqrt{195.1104}}$$

$$f_{11} = \frac{1}{13.9682}$$

$$f_{11} = 0.0716$$

Nilai dari perhitungan normalisasi matriks akan digunakan untuk menentukan solusi ideal positif dan solusi ideal negatif. Hasil perhitungan normalisasi matriks pada seluruh elemen matriks dapat dilihat pada tabel 3.22 dan 3.23:

Tabel 3.22 Hasil Normalisasi Matriks

	F1	F2	F3	F4	F5	F6	F7
S1	0.0716	0.0358	0.0195	0.6443	0.0000	0.7603	0.0055
S2	0.0664	0.0266	0.0102	0.9301	0.0000	0.3586	0.0332
S3	0.0688	0.0413	0.0000	0.4126	0.0000	0.9072	0.0188
S4	0.0333	0.0266	0.0102	0.9311	0.0333	0.3590	0.0333
S5	0.0262	0.0393	0.0000	0.7862	0.0218	0.6157	0.0131
S6	0.0081	0.0243	0.0046	0.2830	0.0047	0.9588	0.0031
S7	0.0425	0.0170	0.0000	0.6798	0.0340	0.7309	0.0189
S8	0.0235	0.0282	0.0061	0.4937	0.0000	0.8688	0.0039
S9	0.0420	0.0168	0.0080	0.4198	0.0000	0.9064	0.0052
S10	0.0794	0.0079	0.0000	0.1587	0.0061	0.9841	0.0041

3.2.4.3 Menentukan Solusi Positif Ideal dan Solusi Negatif Ideal

Solusi positif ideal adalah nilai terbesar dari masing-masing fitur ekstraksi. sedangkan solusi negatif ideal adalah nilai terkecil dari masing-masing fitur ekstraksi. Untuk menghitung solusi positif ideal dan solusi negatif ideal dapat dihitung dengan rumus 2.10 dan 2.11.

$$A^* = \{\max f_{ij} | i = 1,2, \dots, m\}$$
$$= \{f_1{}^*, f_2{}^*, \dots, f_i{}^*, \dots, f_n{}^*\}$$

$$A^- = \{\min f_{ij} | i = 1,2, \dots, m\}$$
$$= \{f_1^-, f_2^-, \dots, f_i^-, \dots, f_n^-\}$$

Keterangan : A^* = nilai maksimal dari proses normalisasi matriks

A^- = nilai minimal dari proses normalisasi matriks

Nilai solusi positif dan negatif ideal akan digunakan pada proses perhitungan metode VIKOR berikutnya yaitu perhitungan nilai untilitas dan regret. Nilai solusi positif dan negatif ideal dapat dilihat pada tabel 3.23.

Tabel 3.23 Solusi Ideal

Solusi ideal	F1	F2	F3	F4	F5	F6	F7
A^*	0.0794	0.0413	0.0195	0.9311	0.0340	0.9841	0.0333
A^-	0.0081	0.0079	0.0000	0.1587	0.0000	0.3586	0.0031

3.2.4.4 Menghitung Utilitas (Si) dan Regret (Ri)

Dalam perhitungan ukuran utilitas dan regret diperlukan nilai bobot fitur (W) yang diberikan oleh seorang pakar yang memiliki latar belakang pendidikan bahasa indonesia yaitu Mahmud Fasya M.A dan Dina Novalia S.S. Bobot fitur yang diberikan dapat dilihat pada tabel 3.21 Pembobotan fitur oleh pakar. Utilitas dan regret dihitung dengan rumus 2.12 dan 2.13.

$$S_i = \sum_{j=1}^{n} W_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)}$$

$$R_i = max_j \left[W_j \frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)} \right]$$

Keterangan : W = bobot dari fitur

f_j^* = nilai maksimal fitur dari normalisasi matriks

f_j^- = nilai minimal fitur dari normalisasi matriks

f_{ij} = nilai fitur dari hasil normalisasi matriks

Berikut adalah contoh perhitungan utiliti (Si) dan regret (Ri) pada nilai normalisasi fitur f_{11} :

$$S_1 = \sum_{j=7}^{1} 0{,}1429\frac{(0.0794 - 0{,}5236)}{(0.0794 - 0.0081)} + 0{,}2143\frac{(0.0413 - 0.0358)}{(0.0413 - 0.0079)}$$
$$+ 0{,}25\frac{(0.0195 - 0.0195)}{(0.0195 - 0)} + 0{,}0714\frac{(0.9311 - 0.6443)}{(0.9311 - 0.1587)}$$
$$+ 0{,}0357\frac{(0.0340 - 0)}{(0.0340 - 0)} + 0.1786\frac{(0.9841 - 0.3102)}{(0.9841 - 0.3586)}$$
$$+ 0{,}1071\frac{(0.0333 - 0.0055)}{(0.0333 - 0.0031)}$$

$$S_1 = \ 0.0156 + 0.0351 + 0 + 0.0265 + 0.0357 + 0.0639 + 0.0986(R_1)$$

$$S_1 = 0{,}2831$$

$$R_1 = max_7\left[W_j\frac{(f_j^* - f_{ij})}{(f_j^* - f_j^-)}\right] = 0{,}1071\frac{(0.0333 - 0.0055)}{(0.0333 - 0.0031)} = 0.0986$$

Nilai dari perhitungan utilitas dan regret akan digunakan dalam tahapan terakhir dalam proses perhitungan metode VIKOR yaitu menghitung indeks VIKOR. nilai utilitas dan regret dari ketiga pakar dapat dilihat pada tabel 3.24, tabel 3.25, dan tabel 3.26. Nilai bobot (W) pakar yang digunakan untuk perhitungan dapat dilihat pada tabel 3.21. berikut adalah hasil perhitungan dengan menggunakan nilai dari pakar 1 oleh Mahmud Fasya M.A.

Tabel 3.24 Utilitas dan Regret dengan Nilai Bobot Pakar 1

Kalimat	Utilitas (Si)	Regret (Ri)
S1	0.2754	0.0986
S2	0.4540	0.1786
S3	0.4284	0.2500
S4	0.4849	0.1785
S5	0.5721	0.2500
S6	0.6481	0.1908
S7	0.6266	0.2500
S8	0.5807	0.1715
S9	0.5846	0.1573
S10	0.6687	0.2500

Kalimat	Utilitas (Si)	Regret (Ri)
Nilai Max	0.6687	0.2500
Nilai Min	0.2754	0.0986

Berikut adalah hasil perhitungan utilitas dan regret dengan menggunakan nilai bobot (W) dari pakar 2 yaitu Dina Novalia S.S. Hasil perhitungan nilai utilitas dan regret pakar 2 dapat dilihat pada tabel 3.25.

Tabel 3.25 Hasil Perhitungan Utilitas dan Regret Nilai Bobot Pakar 2

Kalimat	Utilitas (Si)	Regret (Ri)
S1	0.3178	0.0920
S2	0.4871	0.25
S3	0.3683	0.15
S4	0.5073	0.2498
S5	0.5243	0.15
S6	0.6037	0.15
S7	0.5546	0.15
S8	0.5487	0.1175
S9	0.5542	0.1468
S10	0.5878	0.2
Nilai Max	0.6037	0.25
Nilai Min	0.3178	0.0920

Berikut adalah hasil perhitungan utilitas dan regret dengan menggunakan nilai bobot (W) dari pakar 3 yaitu Tri Ratna Nigrum S,Pd. Hasil perhitungan nilai utilitas dan regret pakar 3 dapat dilihat pada tabel 3.26

Tabel 3.26 Hasil Perhitungan Utilitas dan Regret Nilai Bobot Pakar 3

Kalimat	Utilitas (Si)	Regret (Ri)
S1	0.2728	0.0920
S2	0.3299	0.1431
S3	0.5373	0.3000
S4	0.5154	0.1427
S5	0.6762	0.3000
S6	0.6864	0.2290
S7	0.6107	0.3000
S8	0.6147	0.2058

Kalimat	Utilitas (Si)	Regret (Ri)
S9	0.6204	0.1771
S10	0.6878	0.3000
Nilai Max	0.6878	0.3000
Nilai Min	0.2728	0.0920

3.2.4.5 Menghitung Indeks VIKOR

Indeks VIKOR dihitung dengan menggunakan bobot nilai (v) yang telah ditentukan. Adapun dalam sebuah jurnal menyatakan bahwa nilai (v) yang digunakan adalah yang $\leq 0{,}5$, jika kedua aspek dianggap relevan maka nilai (v) yang digunakan adalah 0,5 [9]. *Utilitas (Si)* dan *Regret (Ri)* akan disubtitusi kedalam persamaan. Berikut adalah contoh perhitungan nilai indeks VIKOR dengan menggunakan rumus 2.14.

$$Q_i = \left[v\ \frac{(S_i - S^-)}{(S^* - S^-)}\right] + \left[(1 - v)\ \frac{(R_i - R^-)}{(R^* - R^-)}\right]$$

Untuk nilai indeks (Q_1) kalimat S1 yang digunakan sebagai contoh perhitungan adalah nilai Utilitas dan Regret dari pakar satu pada tabel 3.24:

$$Q_1 = \left[0{,}5\ \frac{(0.2754 - 0.2754)}{(0.8975 - 0.2754)}\right] + \left[(1 - 0{,}5)\frac{(0.1113 - 0.1113)}{(0{,}25 - 0.1113)}\right] = 0$$

Setelah dilakukan perhitungan indeks vikor akan dilakukan perangkingan dari nilai indeks terbaik adalah nilai indeks terkecil. nilai yang digunakan untuk perhitungan nilai indeks VIKOR untuk pakar 2 adalah nilai Utilitas dan Regret pada tabel 3.24. Hasil perhitungan indeks vikor pakar 1 dapat dilihat pada tabel 3.27.

Tabel 3.27 Hasil Perhitungan Indeks VIKOR untuk Pakar 1

Kalimat	Qi
S1	0
S2	0.4913
S3	0.6945
S4	0.5302
S5	0.8771
S6	0.7785
S7	0.9465

Kalimat	Qi
S8	0.6289
S9	0.5871
S10	1

Dan nilai yang digunakan untuk perhitungan nilai indeks VIKOR untuk pakar 2 adalah nilai Utilitas dan Regret pada tabel 3.25. Hasil perhitungan nilai indeks VIKOR pakar 2 dapat dilihat pada tabel 3.28.

Tabel 3.28 Hasil Perhitungan Indeks VIKOR untuk Pakar 2

Kalimat	Qi
S1	0
S2	0.7961
S3	0.2718
S4	0.8309
S5	0.5446
S6	0.6835
S7	0.5976
S8	0.4846
S9	0.5869
S10	0.8139

Selanjutnya perhitungan indek VIKOR untuk pakar 3 nilai yang digunakan untuk perhitungan nilai indeks VIKOR untuk pakar 3 adalah nilai Utilitas dan Regret pada tabel 3.26. hasil perhitungan nilai indek VIKOR pakar 3 dapat dilihat pada tabel 3.29:

Tabel 3.29 Hasil Perhitungan Indeks VIKOR untuk Pakar 3

Kalimat	Qi
S1	0
S2	0.1915
S3	0.8187
S4	0.4141
S5	0.9860
S6	0.6835
S7	0.8277
S8	0.9071

Kalimat	Qi
S9	0.6234
S10	1

3.2.4.6 Perangkingan dan Hasil Peringkasan

Dalam melakukan perangkingan dari bobot VIKOR adalah dengan mengurutkan nilai Qi (indeks vikor) yang didapatkan dari hasil perhitungan indeks VIKOR mulai dari nilai terkecil sampai ke nilai yang terbesar. Karena alternatif keputusan terbaik adalah nilai indeks vikor terkecil.

Perangkingan dilakukan berdasarkan hasil perhitungan dari indeks VIKOR dari pakar 1 pada tabel 3.27. Hasil perangkingan dari pakar 1 dapat dilihat pada tabel 3.30.

Tabel 3.30 Perangkingan Hasil Bobot VIKOR Pakar 1

Kalimat	Qi	Rangking
S1	0	1
S2	0.4913	2
S4	0.5302	3
S9	0.5871	4
S8	0.6289	5
S3	0.6945	6
S6	0.7785	7
S5	0.8771	8
S7	0.9465	9
S10	1	10

Untuk perangkingan nilai indeks VIKOR pada pakar 2 dilakukan perangkingan berdasarkan nilai pada tabel 3.28. Hasil perangkingan dari pakar 2 dapat dilihat pada tabel 3.31.

Tabel 3.31 Perangkingan Hasil Bobot VIKOR Pakar 2

Kalimat	Qi	Rangking
S1	0	1
S3	0.2718	2
S8	0.4846	3
S5	0.5446	4

Kalimat	Qi	Rangking
S9	0.5869	5
S7	0.5976	6
S6	0.6835	7
S2	0.7961	8
S10	0.8139	9
S4	0.8309	10

Untuk perangkingan nilai indeks VIKOR pada pakar 3 dilakukan perangkingan berdasarkan nilai pada tabel 3.29. Hasil perangkingan dari pakar 3 dapat dilihat pada tabel 3.32.

Tabel 3.32 Perangkingan Hasil Bobot VIKOR Pakar 3

Kalimat	Qi	Rangking
S1	0.0000	1
S2	0.1915	2
S4	0.4141	3
S9	0.6234	4
S8	0.6853	5
S3	0.8187	6
S6	0.8277	7
S7	0.9071	8
S5	0.9860	9
S10	1	10

Dalam pengambilan kalimat yang akan digunakan untuk hasil ringkasan, maka digunakan kompresi 25% dari banyak kalimat yang ada didalam dokumen. Pemilihan hasil ringkasan pakar 1 dengan menggunakan nilai pada tabel 3.30. Hasil ringkasan kalimat oleh pakar 1 dapat dilihat pada tabel 3.33.

Tabel 3.33 Hasil Peringkasan dengan Bobot Fitur Pakar 1

Kalimat	Isi kalimat	Rangking
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.	1
S2	pasalnya pekerja terkena norovirus.	2

Pengambilan hasil ringkasan pakar 2 didapatkan dari hasil perangkingan indek VIKOR pada tabel 3.31. Hasil ringkasan kalimat oleh pakar 2 dapat dilihat pada tabel 3.34.

Tabel 3.34 Hasil Peringkasan dengan Bobot Fitur Pakar 2

Kalimat	Isi kalimat	Rangking
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.	1
S3	virus menyebabkan mual muntah dan diare	2

Pengambilan hasil ringkasan pakar 2 didapatkan dari hasil perangkingan indek VIKOR pada tabel 3.32. Hasil ringkasan kalimat oleh pakar 2 dapat dilihat pada tabel 3.35.

Tabel 3.35 Hasil Peringkasan dengan Bobot Fitur Pakar 3

Kalimat	Isi kalimat	Rangking
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.	1
S2	pasalnya pekerja terkena norovirus.	2

Pada penelitian ini pengelompokan dokumen berdasarkan kepada tema masing-masing yang telah ditentukan sebelumya. Dokumen yang digunakan adalah 4 dokumen untuk masing-masing topik berita. Untuk hasil dari ringkasan masing masing dokumen oleh pakar 1 pada topik pertama yaitu “Norovirus” dapat dilihat pada tabel 3.36.

Tabel 3.36 Detail Jumlah Kalimat Hasil Peringkasan Pakar 1,2 dan 3

	D1	D2	D3	D4	Jumlah
Jml kalimat sebelum	10	10	9	10	39
Jml kalimat sesudah	2	2	2	2	8

Berdasarkan dari hasil pemilihan kalimat yang ditunjukkan oleh tabel 3.36. Hasil kalimat yang dipilih untuk semua pakar dapat dilihat pada tabel 3.37, tabel 3.38, dan tabel 3.39. Kalimat hasil ringkasan ini akan digunakan untuk proses penggabungan yang nantinya akan menjadi hasil ringkasan multi dokumen.

Tabel 3.37 Hasil Ringkasan Dokumen dengan Topik Norovirus Pakar 1

Kalimat	Isi
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.
S2	pasalnya pekerja terkena norovirus.
S3	swiss mengkonfirmasi atletnya terjangkit norovirus olimpiade musim dingin 2018.
S4	norovirus menyebar musim dingin menjangkiti atlet ski swiss berlangsungnya pertandingan.
S5	kemarin total 194 infeksi norovirus dikonfirmasi.
S6	menyebarnya virus bernama norovirus mencoreng citra olimpiade musim dingin pyeongchang 2018.
S7	ancaman mengintai olimpiade musim dingin 2018 pyeongchang.
S8	panitia penyelenggara pocog mengumumkan tingginya norovirus sehari upacara pembukaan jumat 9 februari.

Tabel 3.38 Hasil Ringkasan Dokumen dengan Topik Norovirus Pakar 2

Kalimat	Isi
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.
S2	virus menyebabkan mual muntah dan diare.
S3	membawa anggota tim kondisi pemulihan.
S4	virus menular makanan air orang orang kontak langsung.
S5	47 pasien dikarantina 147 pulih beraktivitas.
S6	norovirus virus menyerang alat pencernaan menyebabkan infeksi mengakibatkan penderita muntah diare sakit perut.
S7	angka penyebaran meningkat tiga kali lipat 42 menjadi 128.
S8	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja bertugas olimpiade musim dingin pyeongchang.

Tabel 3.39 Hasil Ringkasan Dokumen dengan topik Norovirus Pakar 3

Kalimat	Isi
S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.
S2	pasalnya pekerja terkena norovirus.
S3	norovirus menyebar musim dingin menjangkiti atlet ski swiss berlangsungnya pertandingan.

Kalimat	Isi
S4	swiss mengkonfirmasi atletnya terjangkit norovirus olimpiade musim dingin 2018.
S5	kemarin total 194 infeksi norovirus dikonfirmasi.
S6	menyebarnya virus bernama norovirus mencoreng citra olimpiade musim dingin pyeongchang 2018.
S7	ancaman mengintai olimpiade musim dingin 2018 pyeongchang.
S8	panitia penyelenggara pocog mengumumkan tingginya norovirus sehari upacara pembukaan jumat 9 februari.

3.3 Penggabungan Hasil Ringkasan

Setelah hasil ringkasan dari beberapa dokumen telah didapatkan dengan menggunakan metode VIKOR maka keseluruh hasil ringkasan tersebut digabungkan dengan menghitung term kemunculan kata dengan TF-IDF dan menghitung kesamaan antar dokumen dengan *cosine similarity* untuk kemudian disusun dan menghaslilkan ringkasan multi dokumen.

3.3.1 Metode TF-IDF

Metode Tf-Idf adalah metode yang memeriksa dan memberikan bobot hubungan kemunculan kata (*term*) antara dokumen satu dengan dokumen yang lainnya. Metode ini menggabungkan dua konsep untuk perhitungan bobot, yaitu *term frequency* (tf) merupakan frekuensi kemunculan term j pada dokumen i. *Document frequency* (df) adalah banyaknya kalimat dimana suatu term j muncul [18]. Kemunculan kata yang ada pada dokumen dapat dilihat pada tabel 3.35 :

Tabel 3.40 Kemunculan Kata pada Setiap Dokumen

kata	S1	S2	S3	S4	S5	S6	S7	S8	IDF	LOG(N/IDF)
mengumumkan	0	0	0	0	0	0	0	1	1	0.9031
mengintai	0	0	0	0	0	0	1	0	1	0.9031
atlet	0	0	0	1	0	0	0	0	1	0.9031
pertandingan	0	0	0	1	0	0	0	0	1	0.9031
berlangsungnya	0	0	0	1	0	0	0	0	1	0.9031
terkena	0	1	0	0	0	0	0	0	1	0.9031
194	0	0	0	0	1	0	0	0	1	0.9031
pasalnya	0	1	0	0	0	0	0	0	1	0.9031
ski	0	0	0	1	0	0	0	0	1	0.9031

kata	S1	S2	S3	S4	S5	S6	S7	S8	IDF	LOG(N/IDF)
mengisi	1	0	0	0	0	0	0	0	1	0.9031
mengerahkan	1	0	0	0	0	0	0	0	1	0.9031
menjangkiti	0	0	0	1	0	0	0	0	1	0.9031
musim	1	0	1	1	0	1	1	0	5	0.2041
pekerja	1	1	0	0	0	0	0	0	2	0.6021
olimpiade	1	0	1	0	0	1	1	0	4	0.3010
bertugas	1	0	0	0	0	0	0	0	1	0.9031
swiss	0	0	1	1	0	0	0	0	2	0.6021
menyebarnya	0	0	0	0	0	1	0	0	1	0.9031
penyelenggara	0	0	0	0	0	0	0	1	1	0.9031
kemarin	0	0	0	0	1	0	0	0	1	0.9031
infeksi	0	0	0	0	1	0	0	0	1	0.9031
menyebar	0	0	0	1	0	0	0	0	1	0.9031
jumat	0	0	0	0	0	0	0	1	1	0.9031
9	0	0	0	0	0	0	0	1	1	0.9031
mengkonfirmasi	0	0	1	0	0	0	0	0	1	0.9031
dingin	1	0	1	1	0	1	1	0	5	0.2041
ancaman	0	0	0	0	0	0	1	0	1	0.9031
militer	1	0	0	0	0	0	0	0	1	0.9031
sehari	0	0	0	0	0	0	0	1	1	0.9031
pocog	0	0	0	0	0	0	0	1	1	0.9031
pembukaan	0	0	0	0	0	0	0	1	1	0.9031
terjangkit	0	0	1	0	0	0	0	0	1	0.9031
mencoreng	0	0	0	0	0	1	0	0	1	0.9031
dikonfirmasi	0	0	0	0	1	0	0	0	1	0.9031
virus	0	0	0	0	0	1	0	0	1	0.9031
tingginya	0	0	0	0	0	0	0	1	1	0.9031
pyeongchang	2	0	0	0	0	1	1	0	3	0.4260
total	0	0	0	0	1	0	0	0	1	0.9031
norovirus	0	1	1	1	1	1	0	1	6	0.1249
februari	0	0	0	0	0	0	0	1	1	0.9031
citra	0	0	0	0	0	1	0	0	1	0.9031
atletnya	0	0	1	0	0	0	0	0	1	0.9031
panitia	0	0	0	0	0	0	0	1	1	0.9031
bernama	0	0	0	0	0	1	0	0	1	0.9031

kata	S1	S2	S3	S4	S5	S6	S7	S8	IDF	LOG(N/IDF)
korsel	1	0	0	0	0	0	0	0	1	0.9031
pemerintah	1	0	0	0	0	0	0	0	1	0.9031
upacara	0	0	0	0	0	0	0	1	1	0.9031
kekosongan	1	0	0	0	0	0	0	0	1	0.9031
2018	0	0	1	0	0	1	1	0	3	0.4260

Setelah ditemukan kemunculan kata yang ada dalam dokumen.maka selanjutnya dilakukan perhitungan idf . Berikut contoh perhitungan idf pada kata pyeongchang dengan menggunakan rumus 2.15 :

$$idf = log\left(\frac{N}{idf}\right)$$

$$idf = log\left(\frac{8}{3}\right) = 0.579784$$

Keterangan N : jumlah dari seluruh kalimat

Idf : nilai kemunculan kata

Setelah dilakukan perhitungan idf pada tabel 3.34, maka selanjutnya dilakukan perhitungan tf-idf dengan contoh hasil yang dapat dilihat pada tabel 3.36.

Tabel 3.41 Perhitungan TF-IDF

kata	S1	S2	S3	S4	S5	S6	S7	S8
mengumumkan	0	0	0	0	0	0	0	0.9031
mengintai	0	0	0	0	0	0	0.9031	0
atlet	0	0	0	0.9031	0	0	0	0
pertandingan	0	0	0	0.9031	0	0	0	0
berlangsungnya	0	0	0	0.9031	0	0	0	0
terkena	0	0.9031	0	0	0	0	0	0
194	0	0	0	0	0.9031	0	0	0
pasalnya	0	0.9031	0	0	0	0	0	0
ski	0	0	0	0.9031	0	0	0	0
mengisi	0.9031	0	0	0	0	0	0	0
mengerahkan	0.9031	0	0	0	0	0	0	0
menjangkiti	0	0	0	0.9031	0	0	0	0
musim	0.2041	0	0.2041	0.2041	0	0.2041	0.2041	0
pekerja	0.6021	0.6021	0	0	0	0	0	0
olimpiade	0.3010	0	0.3010	0	0	0.3010	0.3010	0

kata	S1	S2	S3	S4	S5	S6	S7	S8
bertugas	0.9031	0	0	0	0	0	0	0
swiss	0	0	0.6021	0.6021	0	0	0	0
menyebarnya	0	0	0	0	0	0.9031	0	0
penyelenggara	0	0	0	0	0	0	0	0.9031
kemarin	0	0	0	0	0.9031	0	0	0
infeksi	0	0	0	0	0.9031	0	0	0
menyebar	0	0	0	0.9031	0	0	0	0
jumat	0	0	0	0	0	0	0	0.9031
9	0	0	0	0	0	0	0	0.9031
mengkonfirmasi	0	0	0.9031	0	0	0	0	0
dingin	0.2041	0	0.2041	0.2041	0	0.2041	0.2041	0
ancaman	0	0	0	0	0	0	0.9031	0
militer	0.9031	0	0	0	0	0	0	0
sehari	0	0	0	0	0	0	0	0.9031
pocog	0	0	0	0	0	0	0	0.9031
pembukaan	0	0	0	0	0	0	0	0.9031
terjangkit	0	0	0.9031	0	0	0	0	0
mencoreng	0	0	0	0	0	0.9031	0	0
dikonfirmasi	0	0	0	0	0.9031	0	0	0
virus	0	0	0	0	0	0.9031	0	0
tingginya	0	0	0	0	0	0	0	0.9031
pyeongchang	0.8519	0	0	0	0	0.4260	0.4260	0
total	0	0	0	0	0.9031	0	0	0
norovirus	0	0.1249	0.1249	0.1249	0.1249	0.1249	0	0.1249
februari	0	0	0	0	0	0	0	0.9031
citra	0	0	0	0	0	0.9031	0	0
atletnya	0	0	0.9031	0	0	0	0	0
panitia	0	0	0	0	0	0	0	0.9031
bernama	0	0	0	0	0	0.9031	0	0
korsel	0.9031	0	0	0	0	0	0	0
pemerintah	0.9031	0	0	0	0	0	0	0
upacara	0.0000	0	0	0	0	0	0	0.9031
kekosongan	0.9031	0	0	0	0	0	0	0
2018	0	0	0.4260	0	0	0.4260	0.4260	0

Berikut adalah contoh perhitungan bobot dari kata “pemerintah” pada dokumen pertama (S1) menggunakan rumus 2.16:

$$Wij = Tf_{i\ j} * idf$$

$$W_{11} = 1 * 0.9031 = 0.9031$$

Keterangan : W_{ij} : bobot kata dari kalimat

Tf_{ij} : jumlah kemunculan kata dalam kalimat

Tabel hasil perhitungan dari TF-IDF akan digunakan pada proses berikutnya yaitu *cosine similarity* yang digunakan untuk mencari nilai kesamaan dari kalimat yang satu dengan yang lain dan kalimat akan disusun berdasarkan nilai tertinggi ke terendah.

3.3.2 *Cosine Similarity*

Cosine similarity merupakan metode untuk mengukur jarak antara dokumen atau kalimat dengan pendekatan relevansi *query*. Semakin besar nilai kesamaan vektor query dengan vektor kalimat maka query tersebut dipandang semakin relevan dengan kalimat [12]. Perhitungan *cosine similarity* dapat dilakukan dengan persamaan 2.17.

$$Sim\ (Si, Sj) = \frac{\sum_{j=1}^{m} W_{ij} * W_{ij}}{\sqrt{\sum_{j=1}^{m} (W_{ij}{}^2)} * \sqrt{\sum_{j=1}^{m} (W_{ij}{}^2)}}$$

Keterangan

Si, Sj : similarity antara kalimat Si dengan kalimat Sj

W_{ij} : bobot kata dari kalimat pada daftar kalimat Si dan Sj

Berikut adalah contoh perhitungan nilai *cosine similarity* antara kalimat S1 dengan kalimat S1 menggunakan rumus 2.17.

$$Sim\ (S1, S1) = \frac{\sum_{j=1}^{m} ((0 * 0) + (* 0) + (0 * 0.9031 \dots (\dots * \dots)}{\sqrt{(0^2) + (0.9031^2) + \cdots (\dots^2)} * \sqrt{(0^2) + (0{,}09031^2) + + \cdots (\dots^2)}}$$

$$= \frac{6.9712}{\sqrt{6.9712}.\sqrt{6.9712}}$$

$$= \frac{6.9712}{2.640 * 2.640}$$

$$= \frac{6.9712}{6.9712} = 1$$

Berikut hasil perhitungan *cosine similarity* seluruh kalimat pada tabel 3.42, tabel 3.43, dan tabel 3.44 untuk nilai bobot dari pakar 1, pakar 2 dan pakar 3 :

Tabel 3.42 Hasil Perhitungan *Cosine Similarity* Pakar 1

	S1	S2	S3	S4	S5	S6	S7	S8
S1	-	0.0919	0.0291	0.0590	0.0136	0	0.2326	0
S2	0.0919	-	0	0	0	0	0	0
S3	0.0291	0	-	0.1477	0.0679	0	0.0529	0
S4	0.0590	0	0.1477	-	0.1380	0	0.1432	0
S5	0.0136	0	0.0679	0.1380	-	0	0.0494	0.1068
S6	0	0	0	0	0	-	0	0
S7	0.2326	0	0.0529	0.1432	0.0494	0	-	0
S8	0	0	0	0	0.1068	0	0	-
Total	0.4263	0.0919	0.2976	0.4880	0.3758	0	0.4781	0.1068

Tabel 3.43 Hasil Perhitungan *Cosine Similarity* Pakar 2

	S1	S2	S3	S4	S5	S6	S7	S8
S1	-	0	0	0	0	0	0	0
S2	0	-	0	0.0194	0	0.0180	0	0
S3	0	0	-	0	0	0	0	0
S4	0	0.0194	0	-	0	0.1682	0	0
S5	0	0	0	0	-	0	0.0703	0
S6	0	0.0180	0	0.1682	0	-	0.0540	0.0603
S7	0	0	0	0	0.0703	0.0540	-	0.0676
S8	0	0	0	0	0	0.0603	0.0676	-
Total	0	0.0374	0	0.1875	0.0703	0.3005	0.1920	0.1279

Tabel 3.44 Hasil Perhitungan *Cosine Similarity* Pakar 3

	S1	S2	S3	S4	S5	S6	S7	S8
S1	-	0.1005	0	0.0205	0	0.0510	0.0772	-
S2	0.1005	-	0	0	0	0	0	0.1005
S3	0	0	-	0.0914	0	0	0	0
S4	0.0205	0	0.0914	-	0	0.0503	0.0760	0.0205
S5	0	0	0	0	-	0	0	0
S6	0.0510	0	0	0.0503	0	-	0.0945	0.0510
S7	0.0772	0	0	0.0760	0	0.0945	-	0.0772

	S1	S2	S3	S4	S5	S6	S7	S8
S8	0	0	0	0	0	0	0	0
Total	0.2492	0.1005	0.0914	0.2382	0	0.1958	0.2476	0.2492

Setelah dilakukan perhitungan *cosine similarity* antar kalimat pada tabel 3.42, tabel 3.43, dan tabel 3.44. Maka nilai dari semua perhitungannya dijumlahkan dan diurutkan dari nilai terbesar ke nilai terkecil. Hasil dari pengurutan kalimat dari seluruh dokumen hasil ringkasan dapat dilihat pada tabel 3.45, tabel 3.46, dan tabel 3.47.

Tabel 3.45 Perangkingan Hasil *Cosine Similarity* Pakar 1

Dokumen	Kalimat	Nilai	Urutan
1	S1	0.2492	1
4	S7	0.2476	2
2	S3	0.2381	3
3	S6	0.1958	4
1	S2	0.1005	5
2	S4	0.0914	6
4	S8	0	7
3	S5	0	8

Tabel 3.46 Perangkingan Hasil *Cosine Similarity* Pakar 2

Dokumen	Kalimat	Nilai	Urutan
1	S2	0.3102	1
3	S6	0.2976	2
4	S8	0.1246	3
2	S4	0.0939	4
4	S7	0.0667	5
3	S5	0	6
2	S3	0	7
1	S1	0	8

Tabel 3.47 Perangkingan Hasil *Cosine Similarity* Pakar 3

Dokumen	Kalimat	Nilai	Urutan
1	S1	0.2492	1
4	S7	0.2476	2

Dokumen	Kalimat	Nilai	Urutan
2	S4	0.1246	3
3	S6	0.1958	4
1	S2	0.1005	5
2	S3	0.0914	6
4	S8	0	7
3	S5	0	8

Dari hasil perangkingan nilai total dari perhitungan *cosine similarity*, maka didapat kalimat hasil peringkasan multi dokumen. Kalimat hasil peringkasan dapat dilihat pada tabel 3.41 dan 3.42.

Tabel 3.48 Kalimat Hasil Ringkasan Multi Dokumen Pakar 1

Dokumen ke	Kalimat	Isi
1	S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.
4	S7	ancaman mengintai olimpiade musim dingin 2018 pyeongchang.
2	S3	swiss mengkonfirmasi atletnya terjangkit norovirus olimpiade musim dingin 2018.
3	S6	menyebarnya virus bernama norovirus mencoreng citra olimpiade musim dingin pyeongchang 2018.
1	S2	pasalnya pekerja terkena norovirus.
2	S4	norovirus menyebar musim dingin menjangkiti atlet ski swiss berlangsungnya pertandingan.
4	S8	panitia penyelenggara pocog mengumumkan tingginya norovirus sehari upacara pembukaan jumat 9 februari.
3	S5	kemarin total 194 infeksi norovirus dikonfirmasi.

Tabel 3.49 Kalimat Hasil Ringkasan Multi Dokumen Pakar 2

Dokumen ke	Kalimat	Isi
1	S2	virus menyebabkan mual muntah diare.
3	S6	norovirus virus menyerang alat pencernaan menyebabkan infeksi mengakibatkan penderita muntah diare sakit perut.
4	S8	pocog terpaksa karantina penyebaran penyakit kontaminasi makanan minuman.
2	S4	virus menular makanan air orang orang kontak langsung.

4	S7	angka penyebaran meningkat tiga kali lipat dari 42 menjadi 128.
3	S5	47 pasien dikarantina 147 pulih beraktivitas.
2	S3	membawa anggota tim kondisi pemulihan.
1	S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja bertugas olimpiade musim dingin pyeongchang.

Tabel 3.50 Kalimat Hasil Ringkasan Multi Dokumen Pakar 3

Dokumen ke	Kalimat	Isi
1	S1	pyeongchang pemerintah korsel mengerahkan militer mengisi kekosongan pekerja akan bertugas olimpiade musim dingin pyeongchang.
4	S7	ancaman mengintai olimpiade musim dingin 2018 pyeongchang.
2	S4	swiss mengkonfirmasi atletnya terjangkit norovirus olimpiade musim dingin 2018.
3	S6	menyebarnya virus bernama norovirus mencoreng citra olimpiade musim dingin pyeongchang 2018.
1	S2	pasalnya pekerja terkena norovirus.
2	S3	norovirus menyebar musim dingin menjangkiti atlet ski swiss berlangsungnya pertandingan.
4	S8	panitia penyelenggara pocog mengumumkan tingginya norovirus sehari upacara pembukaan jumat 9 februari.
3	S5	kemarin total 194 infeksi norovirus dikonfirmasi.

3.4 Analisis Kebutuhan Non-Fungsional

Analisis kebutuhan non-fungsional menjelaskan tentang semua pendukung berjalannya sistem. Kebutuhan tersebut meliputi kebutuhan perangkat keras, kebutuhan perangkat lunak, kebutuhan pengguna yang akan menjalankan sistem yang telah dibuat.

3.4.1 Analisis Kebutuhan Perangkat Keras

Perangkat keras pendukung dari sistem peringkasan otomatis multi dokumen ini adalah sebagai berikut :

Tabel 3.51 Analisis Kebutuhan Perangkat Keras

Perangkat	Spesifikasi
Laptop Asus A46C	Processor intel core i5 1.8Ghz

Perangkat	Spesifikasi
	RAM 4GB
	Keyboard (Internal / Eksternal)
	Mouse / Touch Pad

3.4.2 Analisis Kebutuhan Perangkat Lunak

Perangkat lunak (*software*) pendukung dari sistem peringkasan otomatis multi dokumen ini adalah sebagai berikut :

Tabel 3.52 Analisis Kebutuhan Perangkat Lunak

Software pendukung	Sistem Operasi Windows 10 64-bit
	IntelliJ IDEA V2018.1.2
	JDK Java 8 (64-bit)

3.4.3 Analisis Kebutuhan Pengguna

Untuk menjalankan sistem ini dibutuhkan pengguna (*user*) yang dapat menjalankan fungsi-fungsi yang ada di aplikasi ini dengan baik. Kebutuhan mendasar yang harus dimiliki *user* adalah

Tabel 3.53 Analisis Kebutuhan Pengguna

Pengguna	Keterangan
User	- Dapat mengoperasikan perangkat komputer atau laptop dengan baik - Dapat mengikuti dan mengerti perintah yang diberikan oleh program.

3.5 Analisis Kebutuhan Fungsional

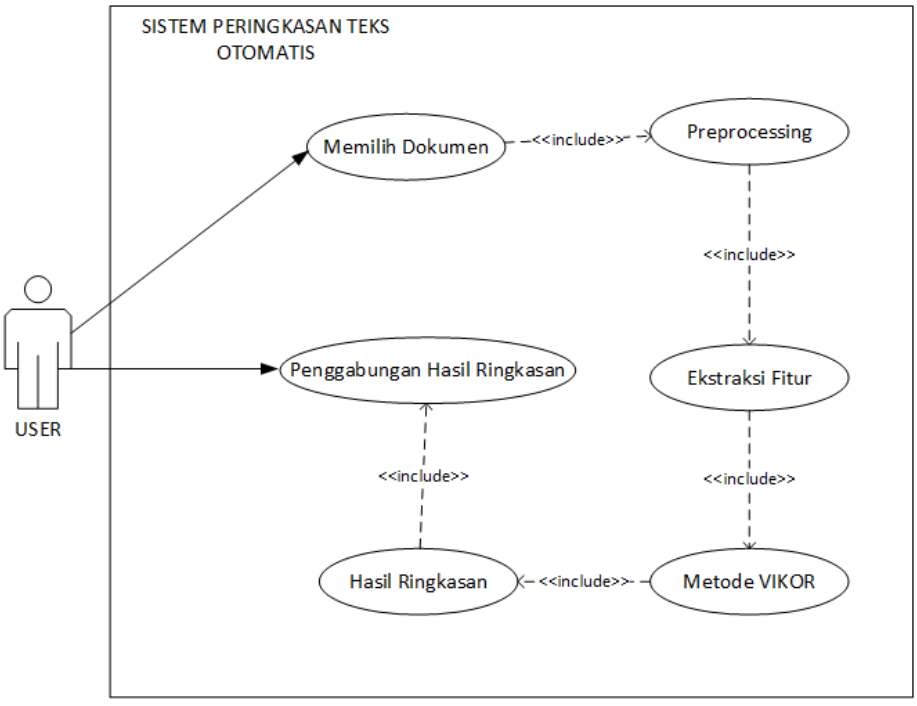
Kebutuhan fungsional disini bertujuan menganalisa proses yang akan diterapkan dalam sistem yang akan dibangun. Analisis yang dilakukan dimodelkan dengan menggunaan UML (*Unified Modelling Language*).

UML digunakan karena merupakan bahasa spesifikasi standar untuk merancang, mengspesifikasikan, dan mendokumentasikan sebuah perangkat lunak berorientasi objek. Bagian-bagian yang digunakan UML untuk menjelaskan analisis dari sistem diantara lain *usecase diagram, usecease skenario, activity diagram, class diagram dan sequence diagram.*

3.5.1 *Use case diagram*

Use case diagram merupakan deskripsi lengkap tentang interaksi yang terjadi antara para actor dengan sistem/perangkat lunak yang sedang dibangun. *Use case diagram* digunakan utnuk menggambarkan fungsi sistem secara umum berdasarkan fungsionalitas sistem dari sudut pandang user. Tujuan dari *use case diagram* adalah untuk menggambarkan dan mengilustrasikan apa yang dapat dilakukan oleh sistem.

Use case diagram dibentuk dari skenario tentang kegunaan sistem yang dinotasikan dengan sebuah *use case*. Setiap skenario menjelaskan suatu alur kegiatan yang dapat diinisialisasi oleh pengguna sistem yang disebut aktor. Pengguna dapat mengamati *use case diagram* untuk dapat memahami utuh tentang pembangunan aplikasi. *Use case* dapat dilihat pada gambar 3.8:



Gambar 3.9 *Usecase Diagram*

3.5.2 Skenario *Use case*

Skenario *use case* menjelaskan tentang deskripsi dari *use case* dan aktor. Skenario *use case* bertujuan untuk menggambarkan interaksi sistem dari setiap *use case* yang digambarkan. Berikut merupakan deskripsi mengenai aktor dapat dilihat pada tabel 3.54 dan deskripsi *usecase* dapat dilihat pada tabel 3.55 :

Tabel 3.54 Deskripsi Aktor

No	Aktor	Deskripsi
1.	Pengguna (*user*)	- Pengguna memilih topik berita yang ada di database - Pengguna memilih dokumen yang akan diproses - Memulai peringkasan

Tabel 3.55 Deskripsi *Use case*

No	Usecase	Deskripsi
1	Memilih topik dan dokumen	Fungsi yang digunakan untuk memilih topik berita dan memilih dokumen yang akan diringkas
2	Mulai peringkasan	Fungsi yang digunakan untuk memulai proses peringkasan.
3	Preprocessing	Fungsi yang digunakan untuk memproses dokumen yang nantinya akan digunakan pada proses ekstraksi fitur.
4	Ekstraksi Fitur	Fungsi yang digunakan untuk mengekstrak dokumen dengan fitur-fitur yang telah ditentukan.
5	Perhitungan Metode VIKOR	Fungsi yang digunakan untuk melakukan perhitungan metode VIKOR.
6	Hasil Ringkasan	Fungsi yang digunakan untuk menampilkan hasil ringkasan dari masing-masing dokumen yang berdasarkan hasil perhitungan metode VIKOR.
7	Penggabungan Hasil Ringkasan	Fungsi yang digunakan untuk menggabungkan hasil ringkasan dan menghasilkan ringkasan multi dokumen.

1. Skenario *Use case* Memilih Dokumen

Tabel 3.56 Skenario *Use case* Memilih Dokumen

Nama Use Case	Memilih dokumen	
Deskripsi	Peroses ini adalah proses pemilihan topik dan dokumen yang akan diproses	
Actor	Pengguna	
Kondisi Awal	Pengguna sudah menyiapkan databese berita	
Skenario utama	No	Reaksi sistem
	1	Sistem menampilkan form.
	2	Pengguna memilih topik berita
	3	Pengguna memilih dokumen yang akan diproses
	4	Sistem menampilkan dokumen yang telah dipilih.
Skenario Alternatif		
	4.1	Dokumen tidak dapat diproses

2. Skenario *Use case* Mulai Peringkasan

Tabel 3.57 Skenario *Use case* Mulai Peringkasan

Nama Use Case	Mulai Peringkasan	
Deskripsi	Peroses ini adalah proses untuk memulai proses peringkasan	
Actor	Pengguna	
Kondisi Awal	Pengguna sudah memilih dokumen berita	
Skenario utama	No	Reaksi sistem
	1	Sistem menampilkan form.
	2	Pengguna memilih topik berita
	3	Pengguna memilih dokumen yang akan diproses
	4	Pengguna menekan tombol MULAI
	5	Sistem menampilkan hasil peringkasan.
Skenario Alternatif		
	5.1	Data tidak dapat diproses

3. Skenario *Use case Preprocessing*

Tabel 3.58 Skenario *Use case Preprocessing*

Nama Use Case	*Preprocessing*	
Deskripsi	Peroses ini adalah untuk memproses data berita yang telah dipilih oleh pengguna	
Actor	Pengguna	
Kondisi Awal	Dokumen berita berhasil diproses oleh sistem	
Skenario utama	No	Reaksi sistem
	1	Sistem mengakses data berita
	2	Sistem melakukan *Preprocessing*
	3	Sistem menampilkan hasil dari *Preprocessing*
Skenario Alternatif		
	3.1	Data berita tidak dapat di-*Preprocessing*

4. Skenario *Use case* Ekstraksi Fitur

Tabel 3.59 Skenario *Use case* Ekstraksi Fitur

Nama Use Case	Ekstraksi Fitur	
Deskripsi	Peroses ini adalah untuk mengekstrak data berita yang telah melalui tahap *preprocessing* dengan menggunakan fitur-fitur ekstraksi	
Actor	Pengguna	
Kondisi Awal	Data berita berhasil di *preprocessing*	
Skenario utama	No	Reaksi sistem
	1	Sistem menerima data berita yang telah melalui tahap *preprocessing*
	2	Sistem melakukan ekstraksi fitur
	3	Sistem menampilkan hasil dari ekstraksi fitur
Skenario Alternatif		
	3.1	Data berita tidak dapat diekstrak

5. Skenario *Use case* Perhitungan Metode VIKOR

Tabel 3.60 Skenario *Use case* Perhitungan Metode VIKOR

Nama Use Case	Perhitungan Metode VIKOR	
Deskripsi	Peroses ini adalah untuk proses data berita yang telah melalui tahap ekstraksi fitur dan mendapatkan hasil ringkasan	
Actor	Pengguna	
Kondisi Awal	Data berita berhasil diekstrak	
Skenario utama	No	Reaksi sistem
	1	Sistem menerima data berita yang telah melalui tahap ekstraksi fitu
	2	Sistem melakukan proses perhitungan metode VIKOR
	3	Sistem menampilkan hasil dari perhitungan metode VIKOR
Skenario Alternatif		
	3.1	Data berita gagal diproses

6. Skenario *Use case* Hasil Ringkasan

Tabel 3.61 Skenario *Use case* Hasil Ringkasan

Nama Use Case	Hasil Ringkasan	
Deskripsi	Peroses ini adalah untuk menampilkan hasil ringkasan dari data berita yang telah melalui proses perhitungan metode VIKOR dan telah dilakukan perangkingan	
Actor	Pengguna	
Kondisi Awal	Perhitungan metode VIKOR berhasil diproses	
Skenario utama	No	Reaksi sistem
	1	Sistem menerima data berita yang telah melalui tahap perhitungan metode VIKOR dan perangkingan
	2	Sistem menampilkan susunan kalimat hasil ringkasan berdasarkan perangkingan
	3	Sistem menampilkan hasil peringkasan.
Skenario Alternatif		
	3.1	Hasil ringkasan tidak dapat ditampilkan

7. Skenario *Use case* Penggabungan Hasil Ringkasan

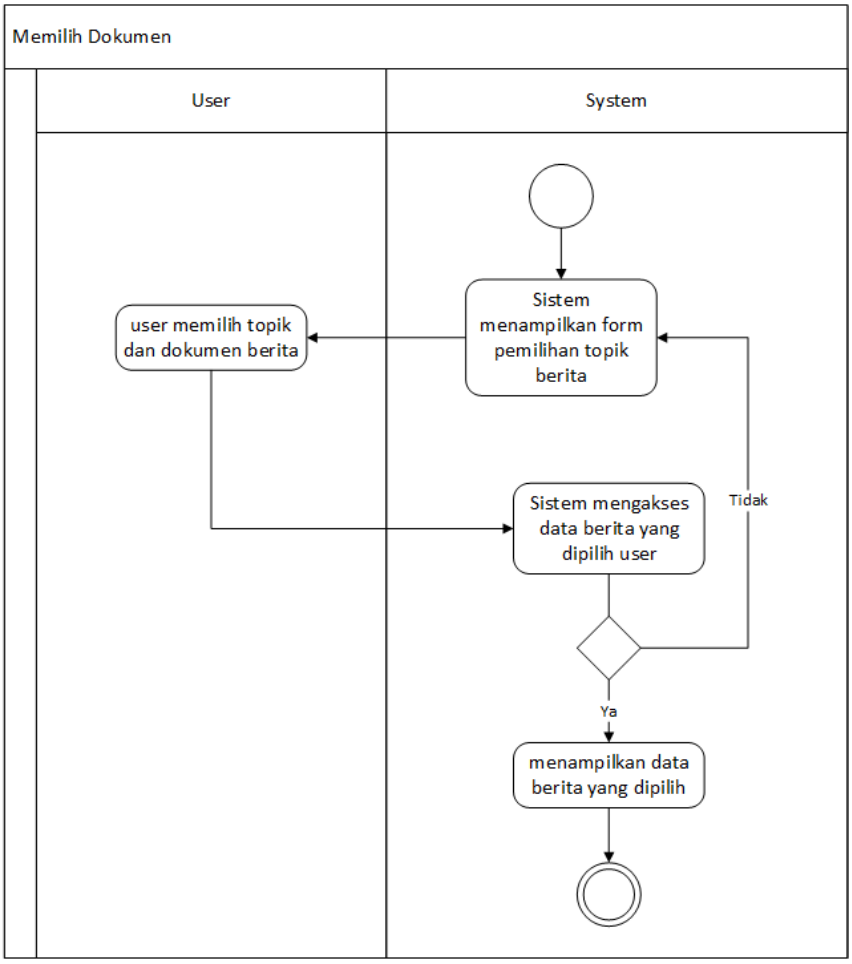
Tabel 3.62 Skenario *Use case* Penggabungan Hasil Ringkasan

Nama Use Case	Hasil Ringkasan	
Deskripsi	Peroses ini adalah untuk menampilkan penggabungan dari hasil ringkasan yang telah didapatkan.	
Actor	Pengguna	
Kondisi Awal	Hasil ringkasan sudah didapatkan	
Skenario utama	No	Reaksi sistem
	1	Sistem menerima data berita yang telah melalui tahap peringkasan
	2	Sistem sistem menggabungkan kalimat hasil ringkasan dari beberapa data berita yang dipilih user
	3	Sistem menampilkan penggabungan hasil ringkasan
Skenario Alternatif		
	3.1	Penggabungan hasil ringkasan gagal diproses

3.5.3 *Activity Diagram*

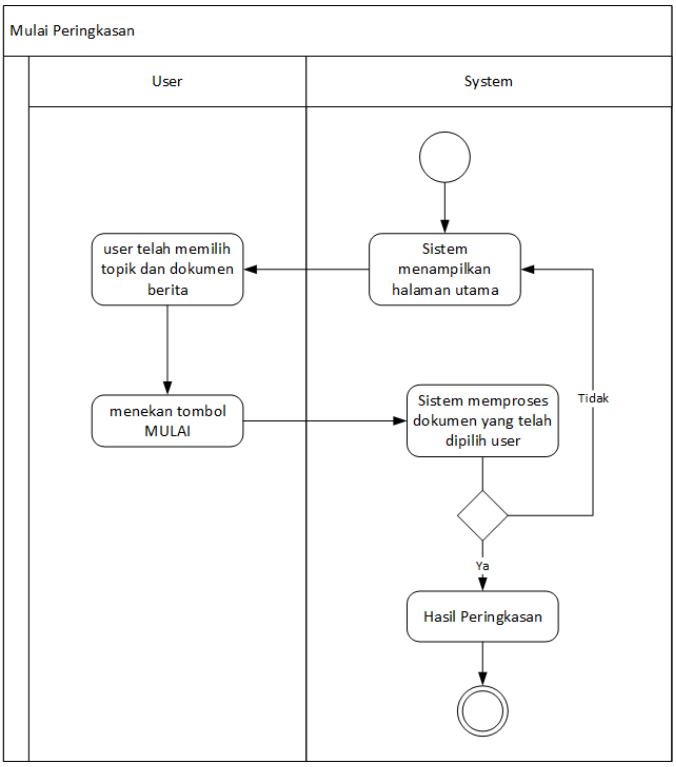
Activity diagram menggambarkan berbagai alur aktivitas dalam sistem yang dirancang, menggambarkan proses dan jalur aktivitas dari level atas secara umum, menggambarkan proses bisnis dan urutan aktivitas dalam sebuah proses, bagaimana masing – masing alur berawal, decision yang mungkin terjadi dan bagaimana alur yang ada berakhir. Berikut merupakan *activity diagram* dari sistem yang dibuat.

1. *Activity Diagram* Memilih Dokumen



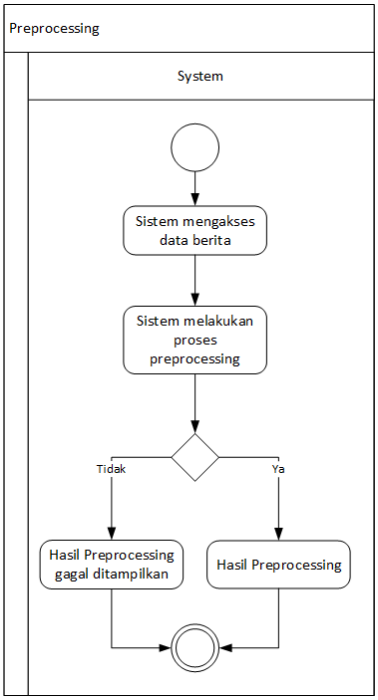
Gambar 3.10 *Activity Diagram* Memilih Dokumen

2. *Activity Diagram* Mulai Peringkasan



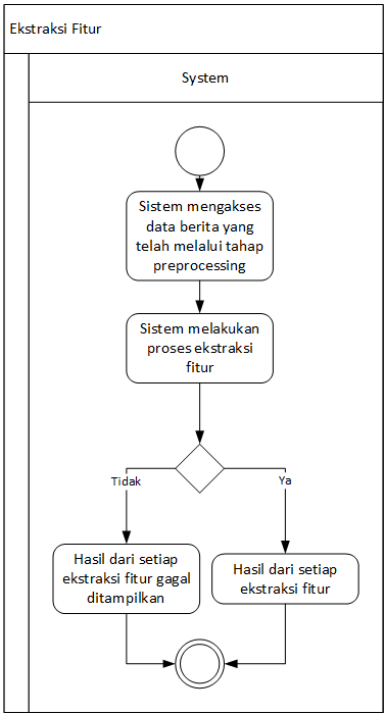
Gambar 3.11 *Activity Diagram* Mulai Peringkasan

3. *Activity Diagram Preprocessing*



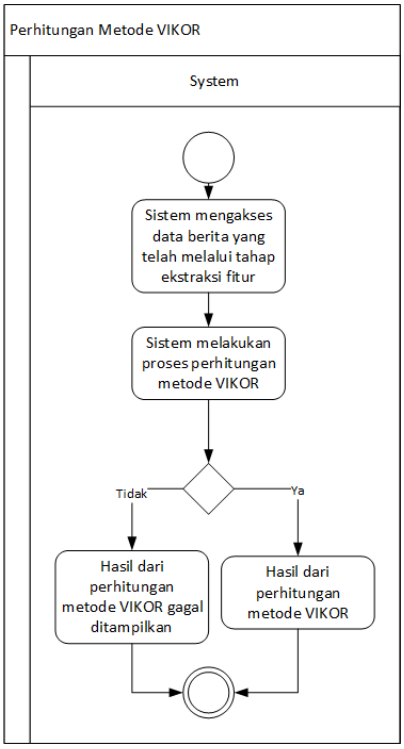
Gambar 3.12 *Activity Diagram Preprocessing*

4. *Activity Diagram* Ekstraksi Fitur



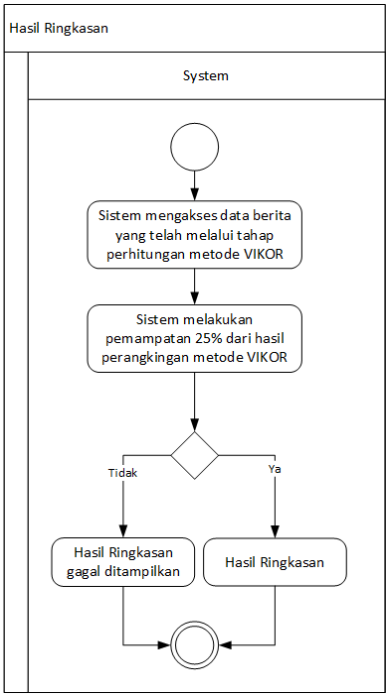
Gambar 3.13 *Activity Diagram* Ekstraksi Fitur

5. *Activity Diagram* Metode Vikor



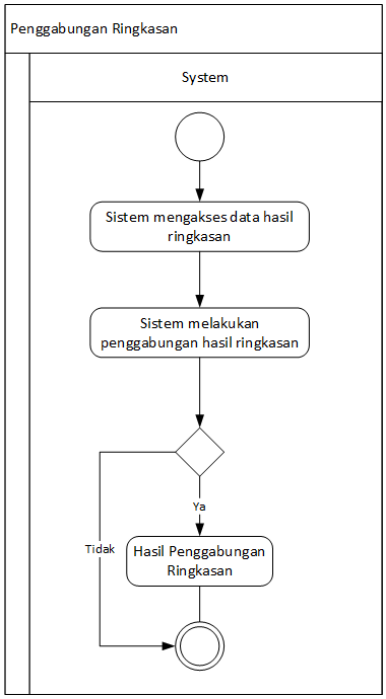
Gambar 3.14 *Activity Diagram* Metode Vikor

6. *Activity Diagram* Hasil Peringkasan



Gambar 3.15 *Activity Diagram* Hasil Peringkasan

7. *Activity Diagram* Penggabungan Hasil Peringkasan

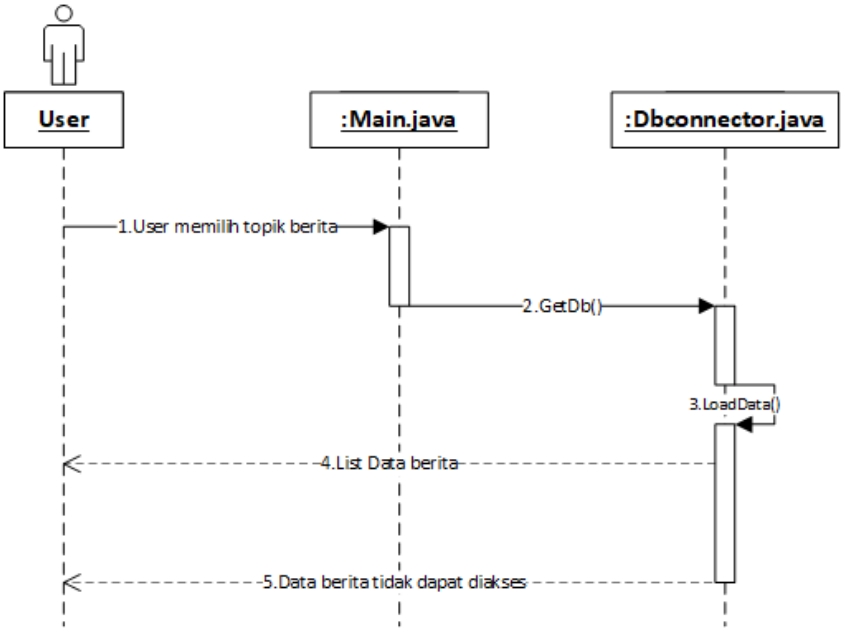


Gambar 3.16 *Activity Diagram* Penggabungan Hasil Peringkasan

3.5.4 *Sequence diagram*

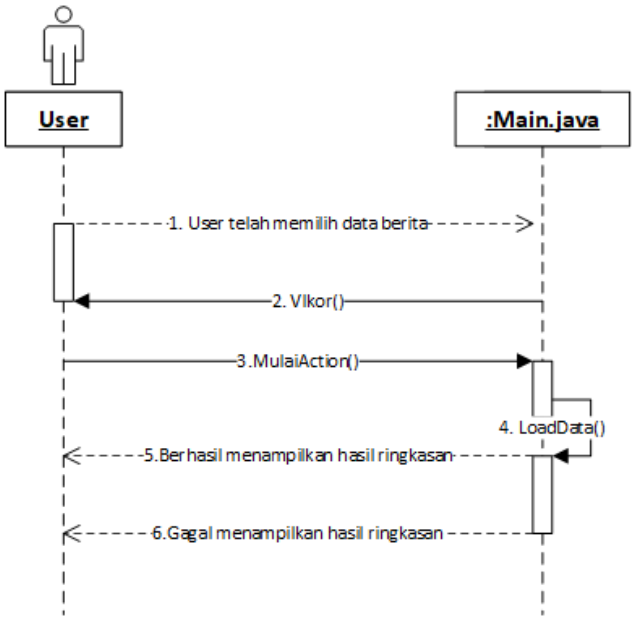
Sequence diagram mendeskripsikan bagaimana entitas dalam sistem berinteraksi. *Sequence diagram* berhubungan erat dengan *use case* diagram, dimana satu *use case* akan menjadi satu *sequence diagram*. Berikut adalah gambaran *sequence diagram*.

1. *Sequence Diagram* Memilih Dokumen



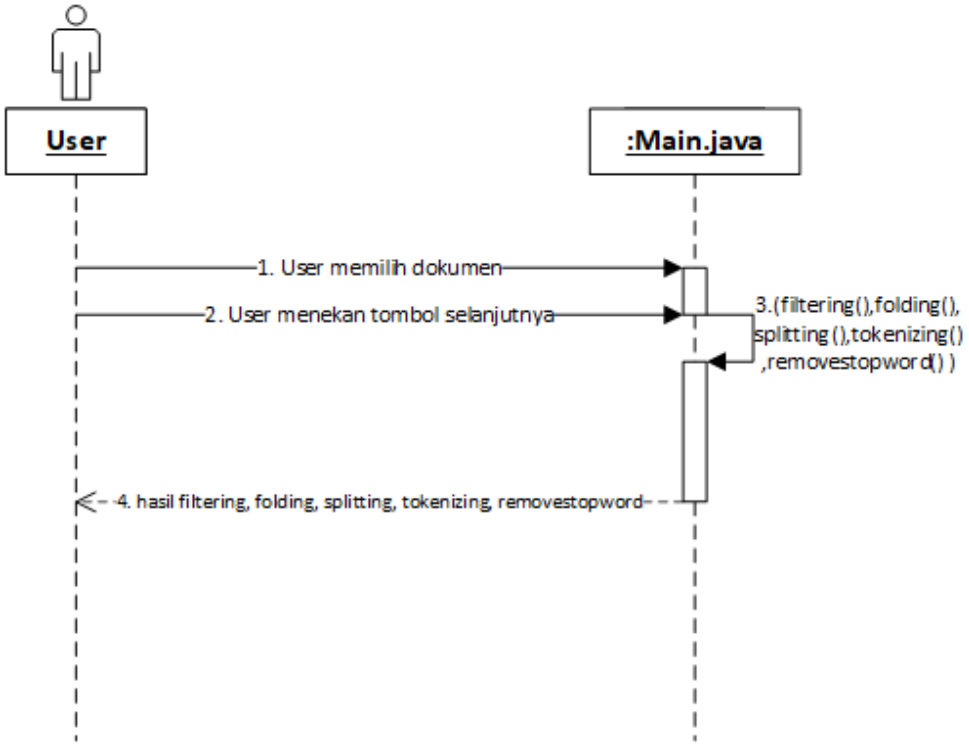
Gambar 3.17 *Sequence Diagram* Memilih Dokumen

2. *Sequence Diagram* Mulai Peringkasan



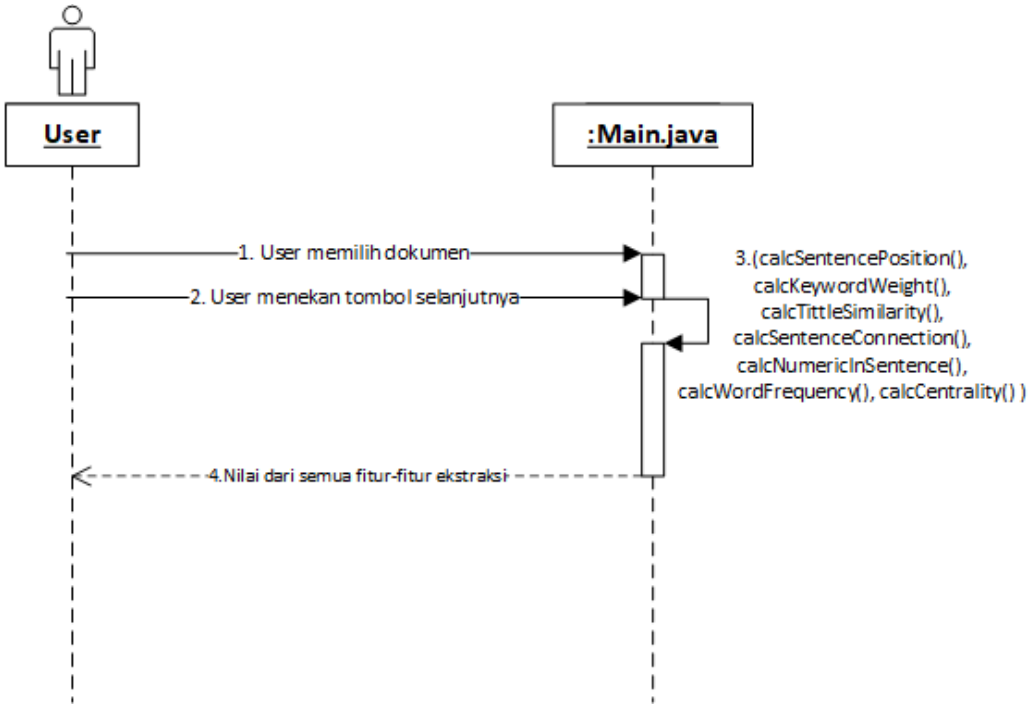
Gambar 3.18 *Sequence Diagram* Mulai Peringkasan

3. *Sequence Diagram Preprocessing*



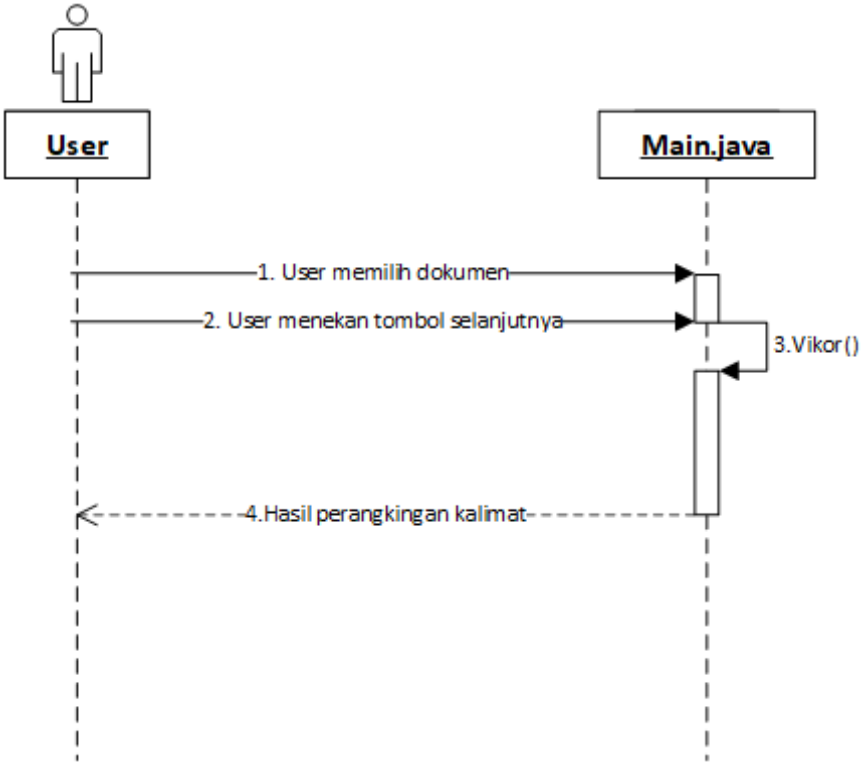
Gambar 3.19 *Sequence Diagram Preprocessing*

4. *Sequence Diagram* Ekstraksi Fitur



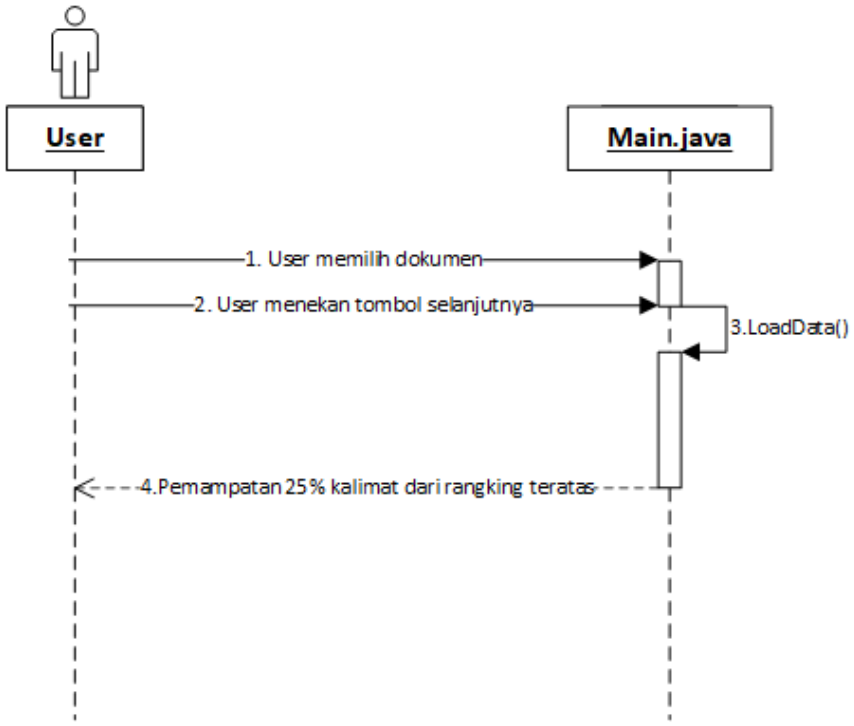
Gambar 3.20 *Sequence Diagram* Ekstraksi Fitur

5. *Sequence Diagram* Metode VIKOR



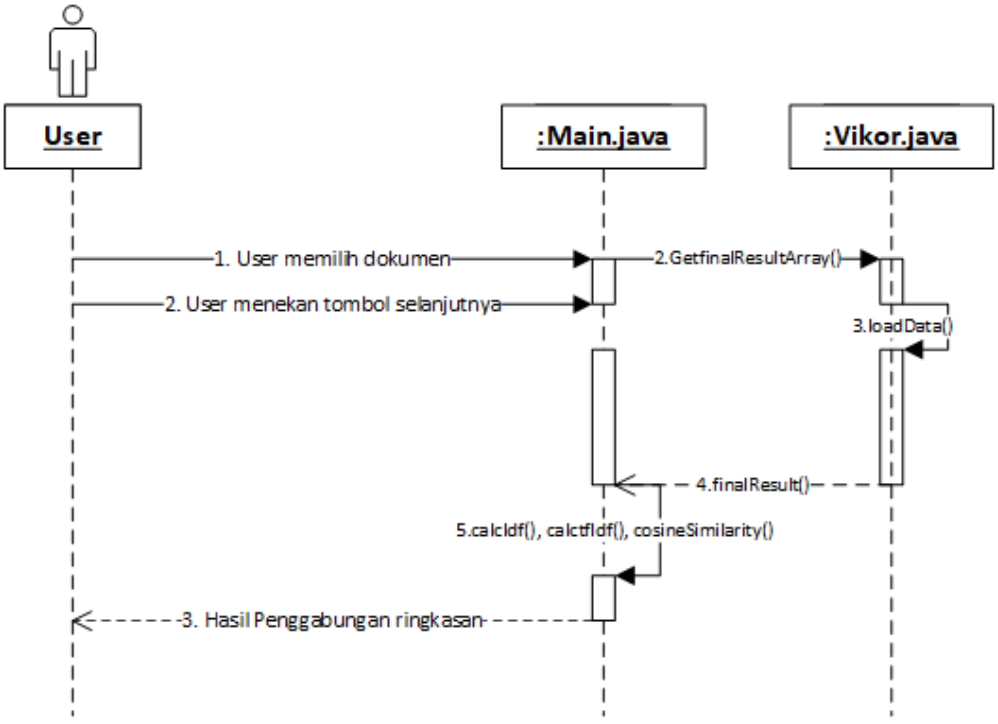
Gambar 3.21 *Sequence Diagram* Metode VIKOR

6. *Sequence Diagram* Hasil Ringkasan



Gambar 3.22 *Sequence Diagram* Hasil Ringkasan

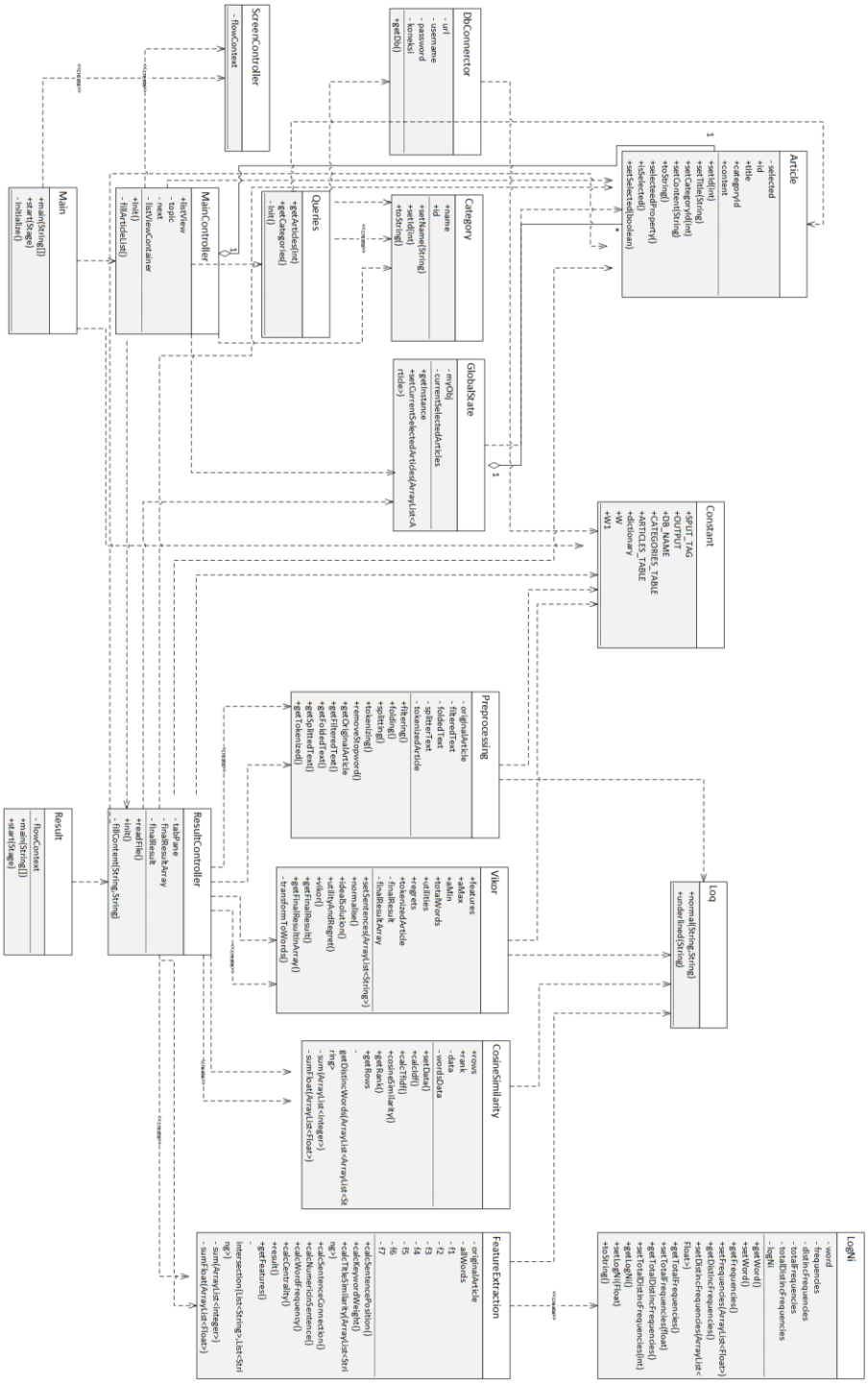
7. *Sequence Diagram* Penggabungan Hasil Ringkasan



Gambar 3.23 *Sequence Diagram* Penggabungan Hasil Ringkasan

3.5.5 *Class Diagram*

Class diagram adalah sebuah spesifikasi yang jika diinstansiasi akan menghasilkan sebuah objek dan merupakan ini dari desain berorientasi objek. Class *diagram* dapat dilihat pada gambar 3.24.



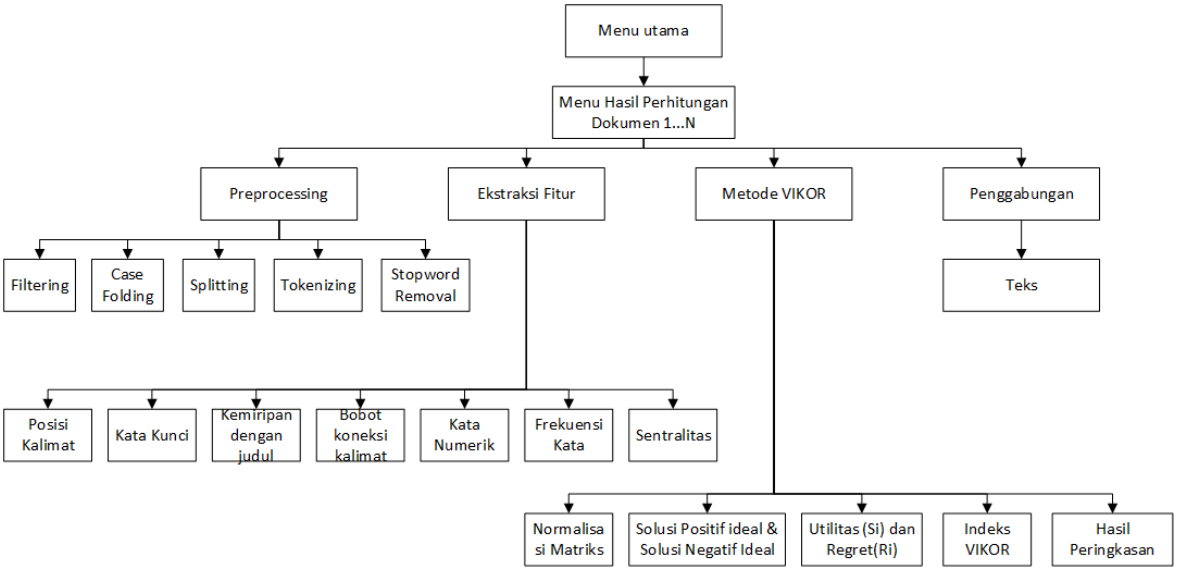
Gambar 3.24 *Class diagram*

3.6 Perancangan Sistem

Perancangan yang digunakan dalam penelitian ini adalah struktur menu, perancangan antarmuka dan perancangan jaringan semantik.

3.6.1 Perancangan Struktur Menu

Perancangan struktur menu yang dibuat adalah membuat sebuah gambaran alur dari pemakaian sistem, sehingga sistem yang dibangun dapat mudah dipahami dan mudah diaplikasikan. Untuk perancangan struktur menu dapat dilihat pada gambar 3.25.



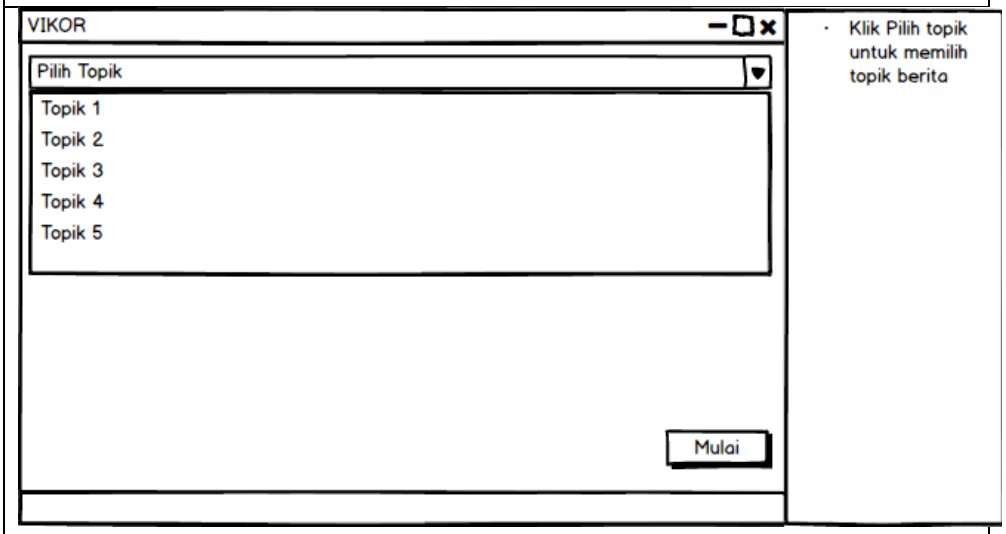
Gambar 25

3.6.2 Perancangan Antarmuka

Perancangan antarmuka dibuat untuk menggambarkan tampilan dari sistem yang nantinya akan menjadi media komunikasi antara pengguna (*user*) dan sistem. Berikut adalah perancangan antar muka sistem yang dibuat.

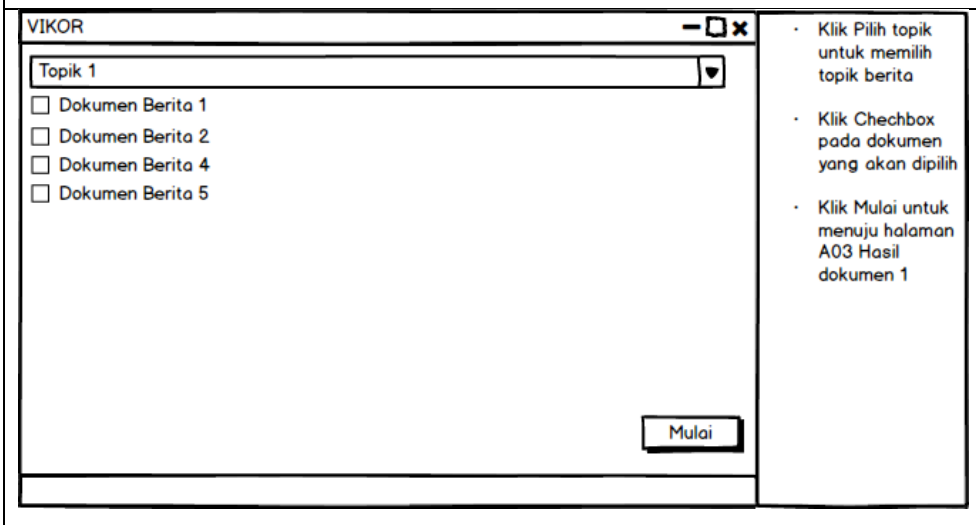
1. **A01 – Halaman utama**

Tabel 3.63 A01 – Halaman utama

A01

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

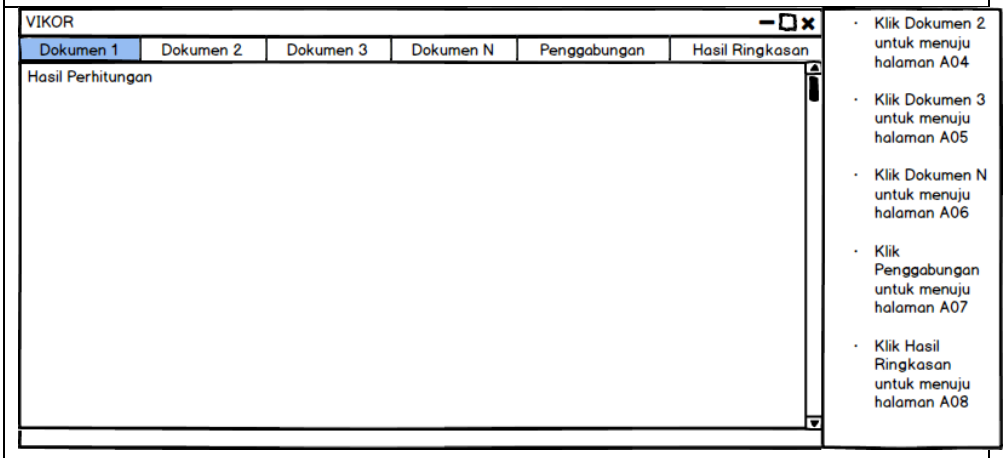
2. **A02 – Halaman utama pilih Dokumen**

Tabel 3.64 A02 – Halaman utama pilih Dokumen

A02

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

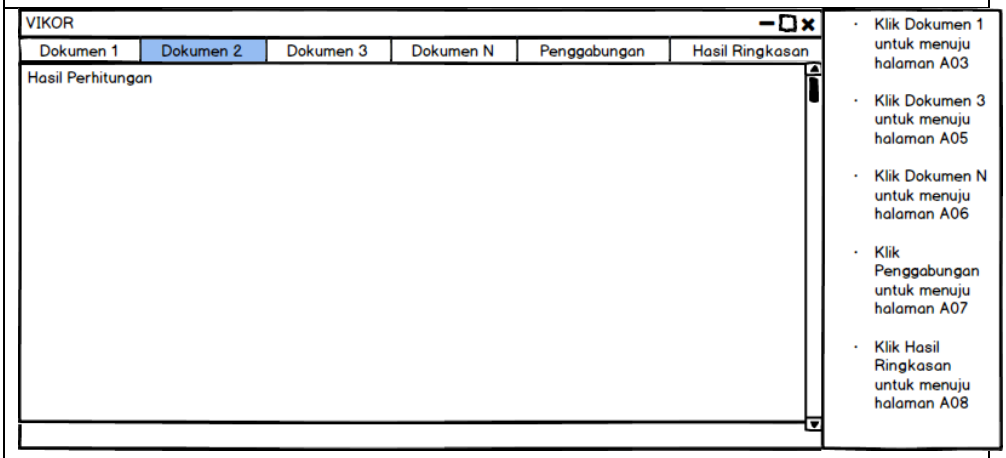
3. A03 – Hasil Dokumen 1

Tabel 3.65 A03 – Hasil Dokumen 1

A03

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

4. A04 – Hasil Dokumen 2

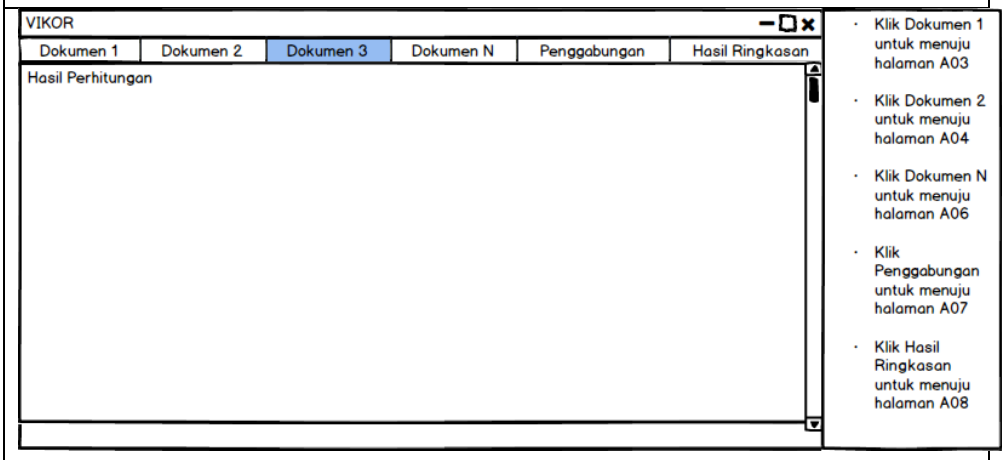
Tabel 3.66 A04 – Hasil Dokumen 2

A04

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

5. A05 – Hasil Dokumen 3

Tabel 3.67 A05 – Hasil Dokumen 3

A05



Font : Arial

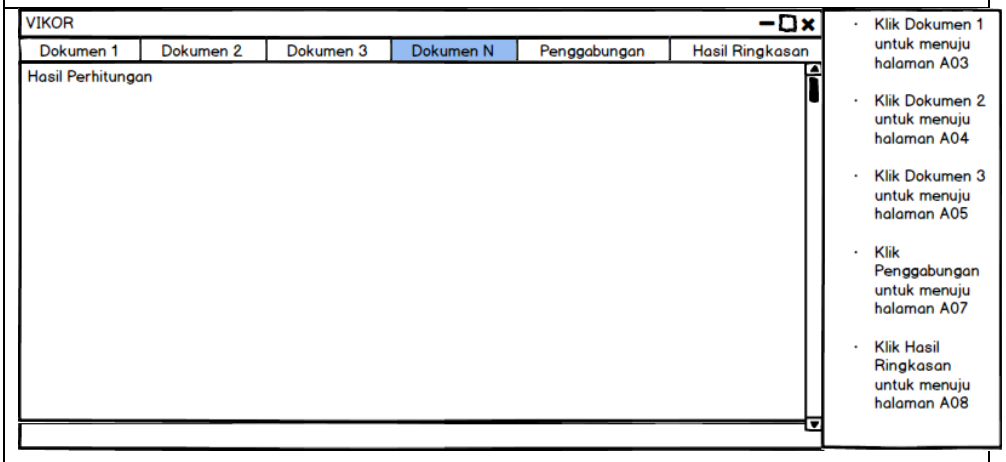
Background Color : disesuaikan

Size Layer : disesuaikan

6. A06 – Hasil Dokumen N

Tabel 3.68 A06 – Hasil Dokumen N

A05



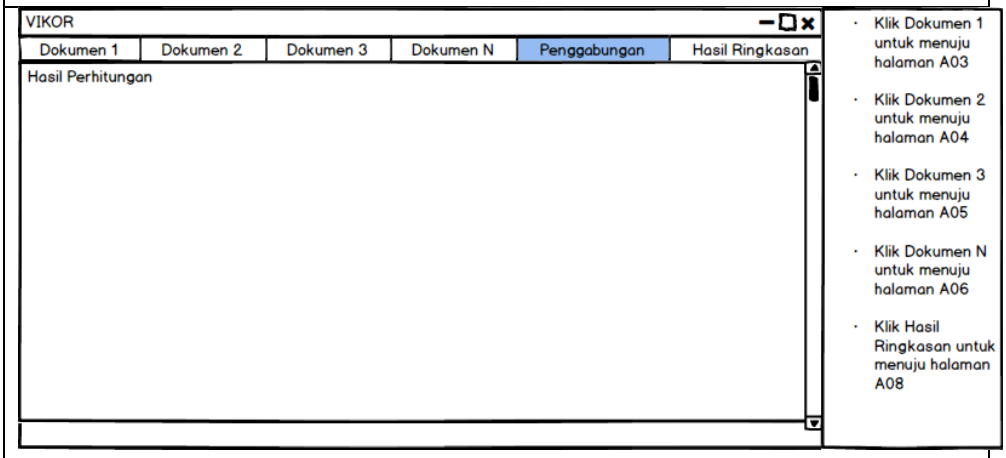
Font : Arial

Background Color : disesuaikan

Size Layer : disesuaikan

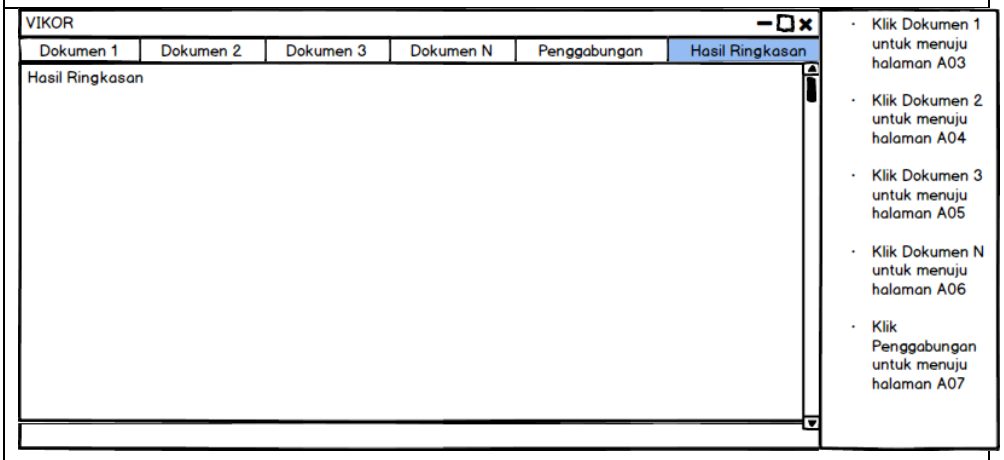
7. A07 – Penggabungan Hasil Ringkasan

Tabel 3.69 A07 – Penggabungan Hasil Ringkasan

A07

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

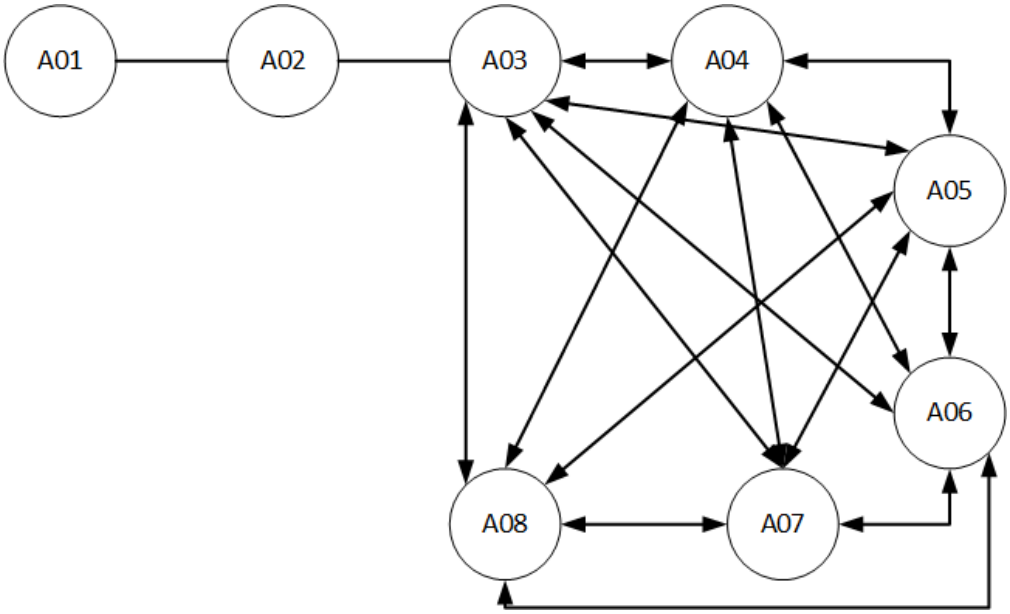
8. A08 – Hasil Ringkasan

Tabel 3.70 A08 – Hasil Ringkasan

A07

Font : Arial Background Color : disesuaikan Size Layer : disesuaikan

3.6.3 Jaringan Semantik

Jaringan semantik adalah sebuah grafis yang menunjukan hubungan antara berbagai objek yang ada pada sistem yang dibuat. Jaringan semantik dapat dilihat pada gambar 3.25.



Gambar 3.26 Jaringan Semantik