BAB 3
ANALISIS DAN PERANCANGAN

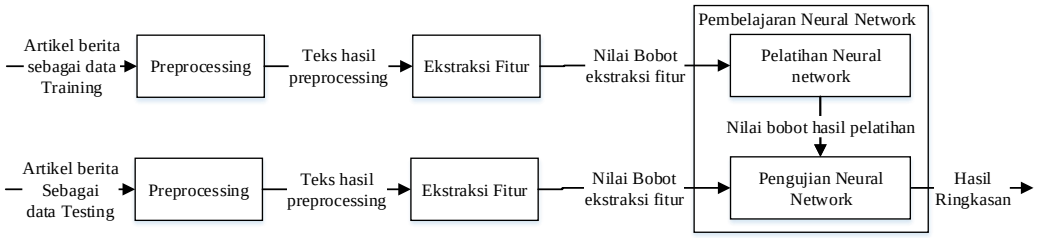
3.1 Analisis Masalah

Penelitian pada peringkasan teks otomatis telah banyak dilakukan terutama pada pengkajian fitur. Salah satunya, penelitian Aristoteles yang mengkaji fitur ekstraksi dengan metode algoritma regresi logistik biner. Berdasarkan hasil pengkajian dengan 20 percobaan yang dilakukan delapan diantaranya mendapatkan akurasi tertinggi yaitu 91.9%. Akan tetapi, dari delapan pengkajian dengan akurasi tertinggi ditemukan bahwa ada lima fitur ekstrasi yang selalu digunakan yaitu posisi kalimat, *positive keyword*, kemiripan kalimat, panjang kalimat dan bobot koneksi antar kalimat [2]. Penelitian ini tidak menggunakan regresi logistik biner melainkan *neural network backpropagation*. Hal ini, berdasarkan hasil penelitian yang membandingkan kedua algoritma tersebut bahwa penggunaan algoritma *neural network backpropagation* memiliki proses pelatihan lebih baik dalam data *training* maupun data *testing* dibandingkan dengan regresi logistik biner [3]. Selain itu, penelitian peringkasan teks otomatis dengan mengimplementasikan algoritma *neural network backpropagation* [4], [5] menunjukkan penggunaan kombinasi fitur ekstraksi yang berbeda ternyata dapat meningkatkan performansi dalam peringkasan. Oleh karena itu, penelitian ini akan dievaluasi tingkat akurasi hasil ringkasan dari penggunaan fitur ekstraksi berdasarkan penemuan dipenelitian sebelumnya [2] serta mengimplementasikan *neural network backpropagation* dalam peringkasan teks otomatis pada artikel Bahasa Indonesia. Penelitian yang dilakukan ini diharapkan menjadi sebuah pendekatan baru dalam penelitian peringkasan teks yang terus dikembangkan sampai saat ini untuk menghasilkan ringkasan teks yang lebih baik.

3.2 Analisis Proses

Analisis proses bertujuan untuk menguraikan permasalahan dan menentukan kebutuhan pada perangkat lunak yang akan dibangun. Analisis proses yang dilakukan meliputi analisis data masukan, analisis *preprocessing*, ekstraksi fitur,

dan analisis metode pembelajaran *neural network*. Berikut adalah *block* diagram dari perancangan peringkasan teks otomatis gambar 3.1.



Gambar 3. 1 Proses Peringkasan Teks

Analisis peringkas teks otomatis memiliki beberapa tahapan. Penjelasan setiap tahap pada peringkasan dapat dilihat sebagai berikut.

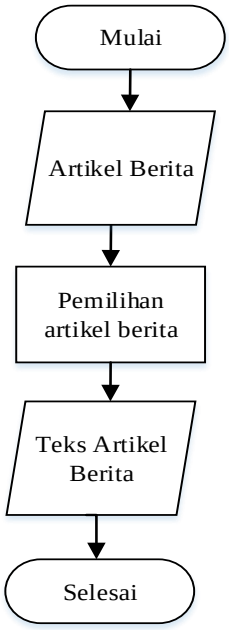
a. Artikel berita sebagai data *training* adalah teks artikel berita yang telah diekstrak menjadi berformat *.txt* sebagai data masukan untuk *preprocessing*. Proses input artikel berita ini sekaligus pemilihan data target output. Data target output adalah data dari pelabelan manual pada artikel yang berisikan nilai 1 dan 0. Data target output ini sebagai data target output pada pembelajaran *neural network backpropagation*.
b. *Preprocessing* merupakan pengelolaan data teks bahan mentah pada proses *tokenizing* kalimat, *casefolding*, *filtering*, *tokenizing* kata dan s*topword removal*.
c. Ekstraksi fitur merupakan tahapan pengelolaan kata pada tiap kalimat setelah melalui tahap *preprocessing*. Fitur ekstraksi yang digunakan yaitu posisi kalimat, *positive keywo*rd, kemiripan antar kalimat, panjang kalimat dan penjumlahan bobot koneksi antar kalimat [2].
d. Artikel berita sebagai data *testing* merupakan artikel dari peringkasan manual yang berada dalam file artikel. Data *testing* diproses melalui *preprocessing* dan ekstraksi fitur yang dimana bobot dari ekstraksi fitur dijadikan data pengujian dalam *neural network* sebagai pembelajaran serta bobot disimpan bersama dengan bobot hasil pengujian sebagai data pembelajaran dalam pelatihan selanjutnya.
e. Pembelajaran *neural network* merupakan tahapan untuk mendapatkan bobot hasil training yang nantinya akan dilakukan *testing* untuk menentukan hasil

ringkasan. Bobot yang didapatkan setelah ekstraksi fitur merupakan nilai masukan dari *layer input* untuk diproses data pelatihan dan pengujian pada *neural network backpropagation*.

f. Hasil ringkasan teks merupakan kalimat-kalimat hasil persentase paling tinggi dari perhitungan *neural network* dengan tingkat kompresi 30% [2].

3.2.1 Analisis Data Masukan

Analisis data masukan pada peringkasan teks artikel berita berbahasa Indonesia merupakan input data training berupa *single* dokumen untuk menghasilkan ringkasan teks. Arsitektur sistem menunjukkan skema proses dari peringkasan teks yang dilakukan ada gambar 3.1. Tahapan pada proses input data untuk peringkasan teks otomatis pada gambar 3.2 adalah sebagai berikut.



Gambar 3. 2 *Flowchart* Diagram Proses Input Data

User melakukan pemilihan artikel pada artikel berita yang nantinya artikel berita yang dipilih akan dijadikan data *input* untuk diproses pada proses selanjutnya yaitu *preprocessing*. Data masukan untuk *testing* pada penelitian ini didapat dari artikel berita pada media online KOMPAS.com dengan judul berita "Google Resmi Luncurkan Fitur Terbaru". Teks dari artikel berita yang dijadikan data *testing* dapat dilihat pada Tabel 3.1

Tabel 3. 1 Data Masukan

Artikel Berita
MENLO PARK - Fitur ‘Pay with Google’ telah diluncurkan untuk pengguna Android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile. Biasanya, para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar. Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay. Google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya. Saat diluncurkan, Pay with Google dapat digunakan untuk 15 tempat, dengan sekira selusin lebih terdaftar sebagai 'coming soon'. Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero. Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez. Fitur itu diberi preview diam-diam saat konferensi I / O bulan Mei. "Mulai hari ini sekarang Anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat," Pali Bhat, VP manajemen produk Google menulis di blog Google. Membayar berbagai tagihan dengan Google menggunakan fitur Google Payment API secara resmi telah diluncurkan secara global. Fitur tersebut dapat digunakan di aplikasi Android di Chrome.Pihak Google mengatakan bahwa Pay with Google tidak dipungut biaya transaksi ketika menggunakannya. Fitur tersebut juga bekerja dengan asisten Google sehingga pengguna dapat memberi tahu asisten untuk membayar orang-orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi.

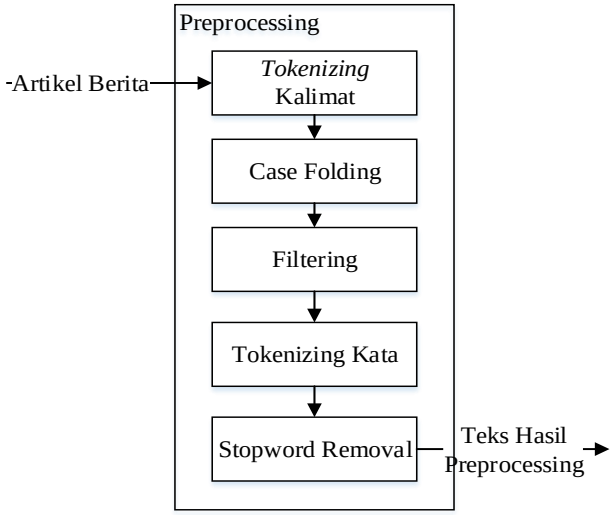
Berdasarkan artikel yang akan digunakan sebagai data set pada proses peringkasan teks otomatis ini, maka terdapat aturan yang harus diperhatikan dalam prosesnya yaitu:

1. Tanda titik yang merupakan tanda sebuah akhir kalimat. Akan tetapi tidak semua titik menunjukkan akhir dari kalimat, dikarenakan tanda titik dapat berada dalam nama gelar, nama alamat, dan lain – lain.
2. Berdasarkan dari data masukan diatas, ditemukan karakter tanda koma, tanda kutip satu atau dua, titik, garing dan strip. Karakter tersebut tidak berpengaruh terhadap data masukan. Oleh karana itu, karakter tersebut akan dihilangkan pada saat proses *filtering*.
3. Karakter selain alfabet yaitu angka atau nomor tidak dihilangkan karena karakter angka atau nomor dalam sebuah kalimat dapat menunjukkan nominal.

4. Berdasarkan data yang telah diamati terdapat sekumpulan kata sering muncul dan tidak memiliki sebuah makna tertentu. Kata tersebut diantaranya kata yang terdapat pada kamus *stopword* yaitu dapat berupa kata penghubung, kata ganti, preposisi, dan lain - lain diantaranya seperti: dia, antara, akan, demi, karena, atau, bahwa, bila, juga, kalau, hingga, bagi, kecuali, oleh, dan lain-lain. Kata tersebut akan dihilangkan para proses stopword removal.

3.2.2 Analisis *Preprocessing*

Analisis *preprocessing* menjelaskan gambaran proses yang diperlukan untuk memilih atau mendapatkan fitur kata yang bersih dan tidak mengandung *noise* yang tidak diperlukan dalam tahap proses selanjutnya. Artikel berita yang telah berformat .txt dijadikan data masukan untuk proses *preprocessing.* Tahapan dari *preprocessing* peringkasan teks yang digunakan terdiri dari *tokenizing* kalimat, *case folding, filtering*, *tokenizing* kata dan *stopword removal*. Alur dari tahapan *preprocessing* dapat dilihat pada gambar 3.3 sebagai berikut.

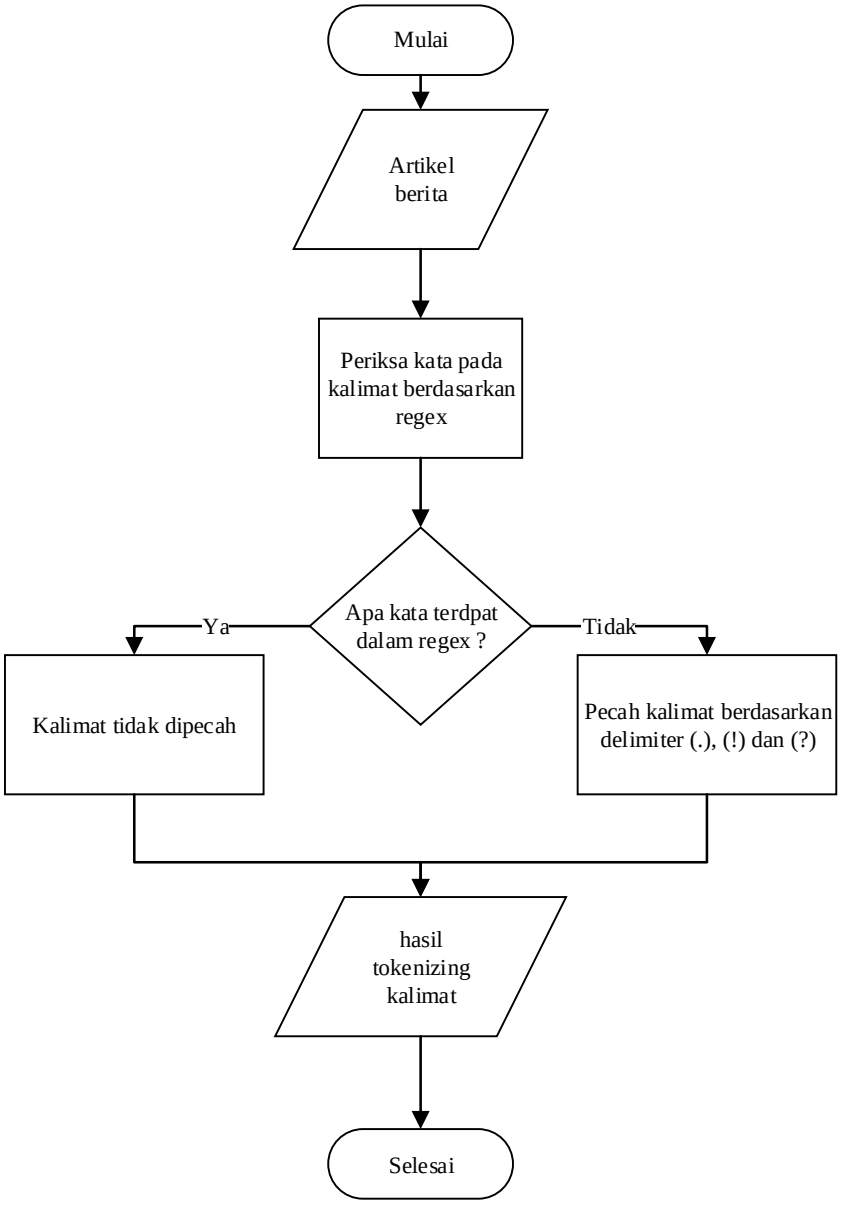


Gambar 3. 3 Proses *Preprocessing*

1. *Tokenizing* Kalimat

Tokenizing kalimat merupakan proses memisahkan kalimat dalam teks dokumen menjadi kumpulan kalimat – kalimat. Proses yang digunakan dalam pemisah kalimat yang berdasarkan pemisah kalimat (*delimiter*) dengan tanda titik "." tanya "?", dan tanda seru "!". Mengatasi masalah *delimiter* "." awal dan akhir kata atau kalimat seperti alamat website, gelar, dan lain – lain yang dapat berpotensi

merusak proses *tokenizing* kalimat. Oleh karena itu, proses ini tidak akan terpisah berdasarkan *delimiter* yang telah ditentukan, contoh yang masuk kedalam daftar pengecualian *regular expression* (*regex*) yaitu: website (.com, .id dll) gelar (dr., ir, s. kom, dll). Gambaran dari *flowchart* diagram *tokenizing* kalimat dapat dilihat pada gambar berikut 3.4 sebagai berikut.



Gambar 3. 4 *Flowchart* Diagram *Tokenizing* kalimat

Kalimat pada artikel berita diperiksa berdasarkan *regex*, jika tidak terdapat dalam *regex* maka proses *tokenizing* kalimat diproses berdasarkan delimiter yang ditentukan dan jika terdapat maka proses *tokenizing* kalimat tidak dilakukan.

Artikel berita sebelum tokenizing kalimat dan hasil dari *tokenizing* kalimat yang merupakan sekumpulan kalimat yang telah terpisah berdasarkan *delimiter*-nya dapat dilihat pada tabel 3.2 sebagai berikut.

Tabel 3. 2 Hasil *Tokenizing* Kalimat

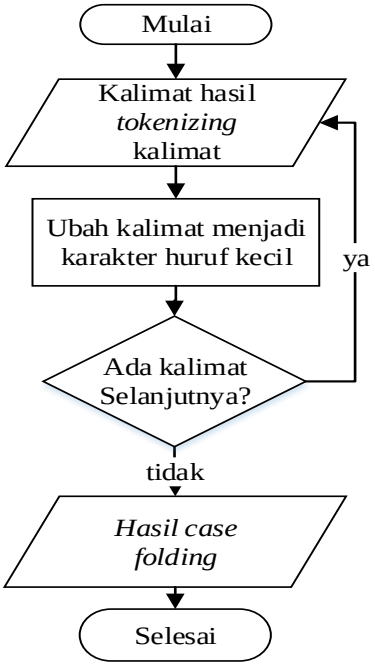
Sebelum	Sesudah	
Artikel Berita	No	Kalimat
MENLO PARK - Fitur ‘Pay with Google’ telah diluncurkan untuk pengguna Android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile. Biasanya, para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar. Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay. Google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya. Saat diluncurkan, Pay with Google dapat digunakan untuk 15 tempat, dengan sekira selusin lebih terdaftar sebagai 'coming soon'. Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero.	S1	MENLO PARK - Fitur Pay with Google telah diluncurkan untuk pengguna Android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile.
	S2	Biasanya, para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar.
	S3	Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay
	S4	Google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya
	S5	Saat diluncurkan, Pay with Google dapat digunakan untuk 15 tempat, dengan sekira selusin lebih terdaftar sebagai 'coming soon'
	S6	Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero

Tabel 3. 2 Hasil *Tokenizing* Kalimat (Lanjutan)

Sebelum	Sesudah	
Artikel Berita	No	Kalimat
Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez. Fitur itu diberi preview diam-diam saat konferensi I / O bulan Mei. "Mulai hari ini sekarang Anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat," Pali Bhat, VP manajemen produk Google menulis di blog Google. Membayar berbagai tagihan dengan Google menggunakan fitur Google Payment API secara resmi telah diluncurkan secara global. Fitur tersebut dapat digunakan di aplikasi Android di Chrome.Pihak Google mengatakan bahwa Pay with Google tidak dipungut biaya transaksi ketika menggunakannya. Fitur tersebut juga bekerja dengan asisten Google sehingga pengguna dapat memberi tahu asisten untuk membayar orang-orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi.	S7	Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez
	S8	Fitur itu diberi preview diam-diam saat konferensi I / O bulan Mei
	S9	Mulai hari ini sekarang Anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat, Pali Bhat, VP manajemen produk Google menulis di blog Google
	S10	Membayar berbagai tagihan dengan Google menggunakan fitur Google Payment API secara resmi telah diluncurkan secara global
	S11	Fitur tersebut dapat digunakan di aplikasi Android di Chrome
	S12	Pihak Google mengatakan bahwa Pay with Google tidak dipungut biaya transaksi ketika menggunakannya
	S13	Fitur tersebut juga bekerja dengan asisten Google sehingga pengguna dapat memberi tahu asisten untuk membayar orang-orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi

2. Case Folding

Case folding merupakan proses mengubah teks menjadi karakter dengan karakter kecil dan hanya karakter huruf ‘a’ sampai ‘z’ yang akan diterima [11]. Gambaran dari *flowchart* diagram *case folding* dapat dilihat pada gambar berikut 3.5 sebagai berikut.



Gambar 3. 5 *Flowchart* Diagram *Case folding*

Data hasil *tokenizing* kalimat dari tiap kalimat diubah menjadi kalimat yang berkarakter huruf kecil sampai kalimat terakhir. Kalimat yang merupakan data hasil tokenizing sebelum proses *case folding* dan sesudah proses *case folding* dapat dilihat pada tabel 3.3 sebagai berikut.

Tabel 3. 3 Hasil *Case Folding*

No	Sebelum	Sesudah
	Kalimat	Kalimat
S1	MENLO PARK - Fitur Pay with Google telah diluncurkan untuk pengguna Android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile.	menlo park fitur pay with google telah diluncurkan untuk pengguna android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile
S2	Biasanya, para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar.	biasanya para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar
S3	Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay	dilansir dari dailymail sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke akun google pengguna melalui google play youtube chrome atau android pay

Tabel 3. 3 Kalimat Sebelum *Case Folding* (Lanjutan)

No	Sebelum	Sesudah
	Kalimat	Kalimat
S5	Saat diluncurkan, Pay with Google dapat digunakan untuk 15 tempat, dengan sekira selusin lebih terdaftar sebagai 'coming soon'	saat diluncurkan pay with google dapat digunakan untuk 15 tempat dengan sekira selusin lebih terdaftar sebagai coming soon
S6	Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero	layanan yang tersedia sekarang termasuk eat24 perusahaan travel planning kayak instacart dan parking reservation service spot hero
S7	Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez	mereka yang terdaftar sebagai coming soon termasuk airbnb perusahaan tiket stubhub papa johns just eat deliveroo dan guzman y gomez
S8	Fitur itu diberi preview diam-diam saat konferensi I / O bulan Mei	fitur itu diberi preview diam diam saat konferensi i o bulan mei
S9	Mulai hari ini sekarang Anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat, Pali Bhat, VP manajemen produk Google menulis di blog Google	mulai hari ini sekarang anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit anda dengan beberapa klik cepat pali bhat vp manajemen produk google menulis di blog google
S10	Membayar berbagai tagihan dengan Google menggunakan fitur Google Payment API secara resmi telah diluncurkan secara global	membayar berbagai tagihan dengan google menggunakan fitur google payment api secara resmi telah diluncurkan secara global
S11	Fitur tersebut dapat digunakan di aplikasi Android di Chrome	fitur tersebut dapat digunakan di aplikasi android di chrome
S12	Pihak Google mengatakan bahwa Pay with Google tidak dipungut biaya transaksi ketika menggunakannya	pihak google mengatakan bahwa pay with google tidak dipungut biaya transaksi ketika menggunakannya
S13	Fitur tersebut juga bekerja dengan asisten Google sehingga pengguna dapat memberi tahu asisten untuk membayar orang-orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi	fitur tersebut juga bekerja dengan asisten google sehingga pengguna dapat memberi tahu asisten untuk membayar orang orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi

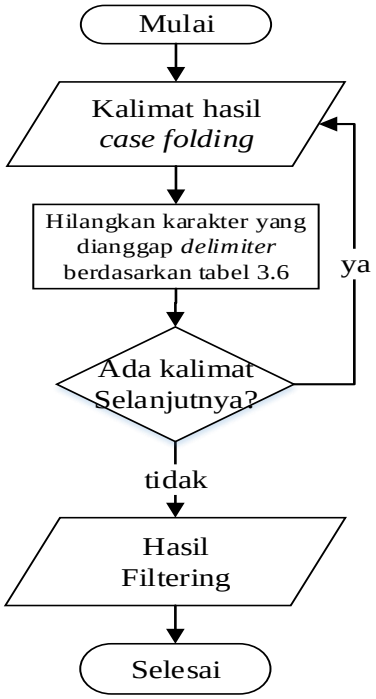
3. *Filtering*

Filtering merupakan proses menghilangkan karakter selain huruf yang dianggap sebagai *delimiter* tidak *valid.* List karakter delimiter yang dihapus dapat dilihat pada tabel 3.4 sebagai berikut.

Tabel 3. 4 List Karakter *Delimiter*

Karakter					
!	&	„	;	]	,
@	*	“	/	|	`
#	(	-	?	\	%
$	)	<	{	+	^
_	.	>	}	[	:
=	,				

Adapun gambaran dari flowchart diagram filtering dapat dilihat pada gambar 3.6 sebagai berikut.



Gambar 3. 6 *Flowchart* Diagram *Filtering*

Data hasil case folding diproses untuk menghilangkan karakter yang terdapat pada tabel 3.4 sehingga kalimat bersih dari karakter yang dianggap sebagai *delimiter*. Kalimat yang merupakan data hasil *case folding* sebelum proses *filtering* dan sesudah proses *filtering* dapat dilihat pada tabel 3.5 sebagai berikut.

Tabel 3. 5 Hasil *Filtering*

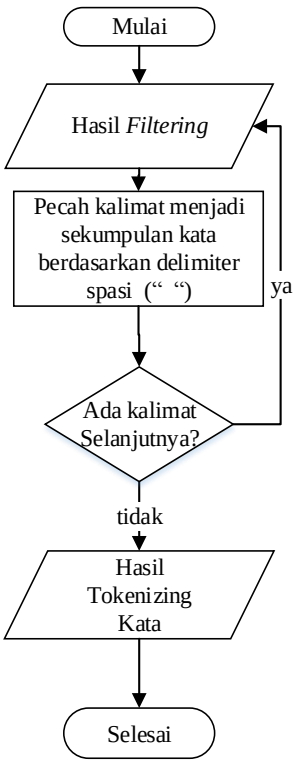
No	Sebelum	Sesudah
	Kalimat	Kalimat
S1	menlo park fitur pay with google telah diluncurkan untuk pengguna android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile	menlo park fitur pay with google telah diluncurkan untuk pengguna android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile
S2	biasanya para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar	biasanya para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar
S3	dilansir dari dailymail sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke akun google pengguna melalui google play youtube chrome atau android pay	dilansir dari dailymail sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke akun google pengguna melalui google play youtube chrome atau android pay
S4	google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya	google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya
S5	saat diluncurkan pay with google dapat digunakan untuk 15 tempat dengan sekira selusin lebih terdaftar sebagai coming soon	saat diluncurkan pay with google dapat digunakan untuk 15 tempat dengan sekira selusin lebih terdaftar sebagai coming soon
S6	layanan yang tersedia sekarang termasuk eat24 perusahaan travel planning kayak instacart dan parking reservation service spot hero	layanan yang tersedia sekarang termasuk eat24 perusahaan travel planning kayak instacart dan parking reservation service spot hero
S7	mereka yang terdaftar sebagai coming soon termasuk airbnb perusahaan tiket stubhub papa johns just eat deliveroo dan guzman y gomez	mereka yang terdaftar sebagai coming soon termasuk airbnb perusahaan tiket stubhub papa johns just eat deliveroo dan guzman y gomez

Tabel 3. 5 Hasil *Filtering* (Lanjutan)

No	Sebelum	Sesudah
	Kalimat	Kalimat
S8	fitur itu diberi preview diam diam saat konferensi i o bulan mei	fitur itu diberi preview diam diam saat konferensi i o bulan mei
S9	mulai hari ini sekarang anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit anda dengan beberapa klik cepat pali bhat vp manajemen produk google menulis di blog google	mulai hari ini sekarang anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit anda dengan beberapa klik cepat pali bhat vp manajemen produk google menulis di blog google
S10	membayar berbagai tagihan dengan google menggunakan fitur google payment api secara resmi telah diluncurkan secara global	membayar berbagai tagihan dengan google menggunakan fitur google payment api secara resmi telah diluncurkan secara global
S11	fitur tersebut dapat digunakan di aplikasi android di chrome	fitur tersebut dapat digunakan di aplikasi android di chrome
S12	pihak google mengatakan bahwa pay with google tidak dipungut biaya transaksi ketika menggunakannya	pihak google mengatakan bahwa pay with google tidak dipungut biaya transaksi ketika menggunakannya
S13	fitur tersebut juga bekerja dengan asisten google sehingga pengguna dapat memberi tahu asisten untuk membayar orang orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi	fitur tersebut juga bekerja dengan asisten google sehingga pengguna dapat memberi tahu asisten untuk membayar orang orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi

4. *Tokenizing* Kata

Tokenizing kata merupakan proses memecah kalimat menjadi kata – kata tunggal. Pemotongan kalimat berdasarkan delimiter yang menyusunnya, yaitu spasi (" "). Gambaran dari *flowchart* diagram *tokenizing* kata dapat dilihat pada gambar 3.7 sebagai berikut.



Gambar 3. 7 *Flowchart* Diagram *Tokenizing* Kata

Data hasil *filtering* yang merupakan sekumpulan kalimat yang telah bersih dari nose kemudian diproses tokenizing kata. Kalimat pada data hasil *filtering* dipecah menjadi kata – kata berdasarkan delimiter spasi (“ “). Proses ini dilakukan hingga kalimat terakhir, sehingga hanya berupa kata – kata berdasarkan posisi kalimat pada dokumen. Kalimat yang merupakan data hasil *filtering* sebelum proses *tokenizing* kata dan sesudah *tokenizing* kata dapat dilihat pada tabel 3.6 sebagai berikut.

Tabel 3. 6 Hasil Tokenizing Kata

	Sebelum	Sesudah					
No	Kalimat	Kata					
S1	menlo park fitur pay with google telah diluncurkan untuk pengguna android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile	menlo	google	android	dapat	pada	
		park	telah	di	mempercepat	perangkat	
		fitur	diluncurkan	seluruh	proses	mobile	
		pay	untuk	dunia	checkout		
		with	pengguna	agar	online		
S2	biasanya para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar	biasanya	beberapa	situs			
		para	harus	formulir	masuk	langganan	
		pengguna	mengisi	untuk	berbagai	berbayar	
S3	dilansir dari dailymail sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke akun google pengguna melalui google play youtube chrome atau android pay	dilansir	pengguna	kartu	dapat	pengguna	chrome
		Dari	dapat	kredit	ditambahkan	melalui	atau
		dailymail	memiliki	atau	ke	google	android
		sekarang	salah	debit	akun	play	pay
		para	1	yang	google	youtube	chrome
S4	google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya	google	pembayaran	layanan	mereka	mengetikkannya	
		nantinya	dan	dengan	yang		
		akan	informasi	informasi	berarti		
		mengirimkan	pengiriman	dari	tidak		
		informasi	pengguna	akun	perlu		
S5	saat diluncurkan pay with google dapat digunakan untuk 15 tempat dengan sekira selusin lebih terdaftar sebagai coming soon	saat	with	tempat	untuk	selusin	sebagai
		diluncurkan	google	dapat	dengan	lebih	coming
		pay	15	digunakan	sekira	terdaftar	soon

Tabel 3. 6 Hasil Tokenizing Kata (Lanjutan)

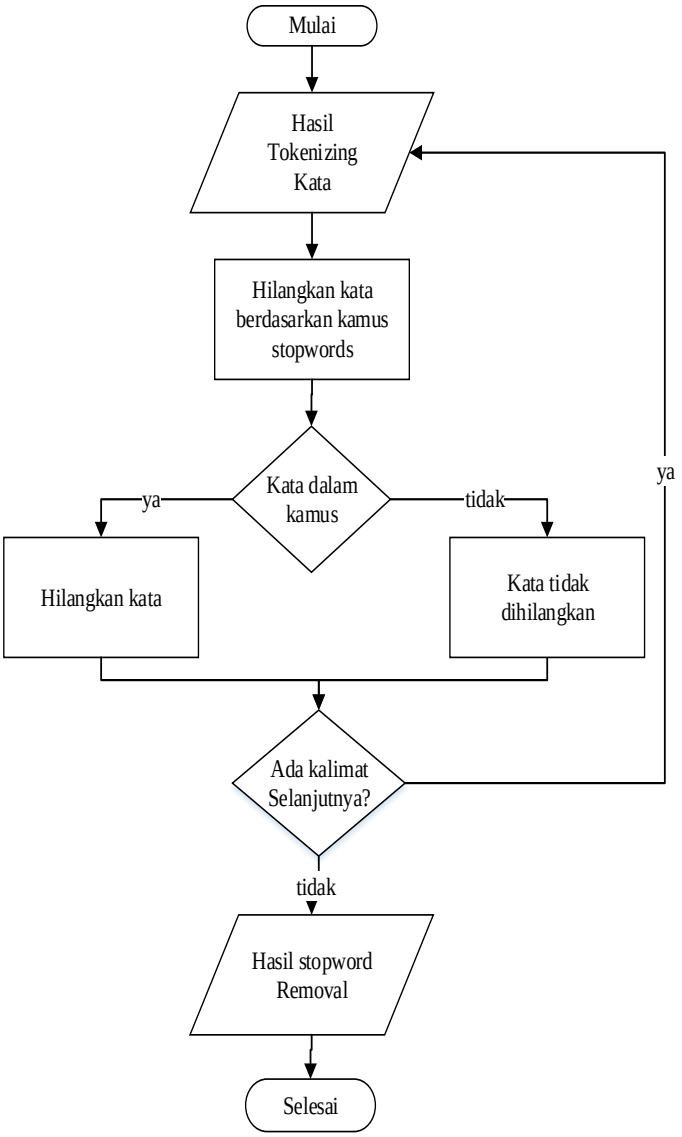
	Sebelum	Sesudah						
No	Kalimat	Kata						
S6	layanan yang tersedia sekarang termasuk eat24 perusahaan travel planning kayak instacart dan parking reservation service spot hero	layanan	eat24	instacart	spot			
		yang	perusahaan	dan	hero			
		tersedia	travel	parking				
		sekarang	planning	reservation				
		termasuk	kayak	service				
S7	mereka yang terdaftar sebagai coming soon termasuk airbnb perusahaan tiket stubhub papa johns just eat deliveroo dan guzman y gomez	mereka	soon	stubhub	deliveroo			
		yang	termasuk	papa	dan			
		terdaftar	airbnb	johns	guzman			
		sebagai	perusahaan	just	y			
		coming	tiket	eat	gomez			
S8	fitur itu diberi preview diam diam saat konferensi i o bulan mei	fitur	diam	bulan				
		itu	saat	mei				
		diberi	konferensi					
		preview	i					
		diam	o					
S9	mulai hari ini sekarang anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit anda dengan beberapa klik cepat pali bhat vp manajemen produk google menulis di blog google	mulai	dapat	banyak	favorit	cepat	produk	google
		hari	mempercepat	aplikasi	anda	pali	google	
		ini	checkout	dan	dengan	bhat	menulis	
		sekarang	online	situs	beberapa	vp	di	
		anda	di	web	klik	manajemen	blog	

Tabel 3. 6 Hasil Tokenizing Kata (Lanjutan)

No	Sebelum	Sesudah					
	Kalimat	Kata					
S10	pihak google mengatakan bahwa pay with google tidak dipungut biaya transaksi ketika menggunakannya	pihak	google	ketika			
		mengatakan	tidak	menggunakannya			
		bahwa	dipungut				
		pay	biaya				
		with	transaksi				
S11	fitur tersebut juga bekerja dengan asisten google sehingga pengguna dapat memberi tahu asisten untuk membayar orang orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi	fitur	asisten	memberi	orang	teman	meminta
		tersebut	google	tahu	orang	pengguna	otentikasi
		juga	sehingga	asisten	tertentu	dan	
		bekerja	pengguna	untuk	dari	kemudian	
		dengan	dapat	membayar	daftar	akan	
S12	membayar berbagai tagihan dengan google menggunakan fitur google payment api secara resmi telah diluncurkan secara global	membayar	menggunakan	secara	global		
		berbagai	fitur	resmi			
		tagihan	google	telah			
		dengan	payment	diluncurkan			
		goolge	api	secara			
S13	fitur tersebut dapat digunakan di aplikasi android di chrome	fitur	dapat	di	android	chrome	
		tersebut	digunakan	aplikasi	di	google	

5. Stopword Removal

Stopword removal merupakan kata-kata yang tidak memiliki makna atau kata yang kurang berarti dan sering muncul dalam kumpulan kata-kata [11]. Gambaran dari *flowchart* diagram *stopword removal* dapat dilihat pada gambar 3.7 sebagai berikut.



Gambar 3. 8 *Flowchart* Diagram *Stopword Removal*

Data hasil tokenizing kata merupakan data masukan pada proses *stopword* removal. Kata pada data tokenizing kata diperiksa berdasarkan kamus yang berada pada *stopword*. Sekumpulan kata yang termasuk dalam kata s*topword* dapat berupa kata penghubung, kata ganti, preposisi, dan lain - lain diantaranya seperti: dia,

antara, akan, demi, karena, atau, bahwa, bila, juga, kalau, hingga, bagi, kecuali, oleh, dan lain-lain. Penghapusan kata yang merupakan termasuk atau terdapat pada kamus akan dihilangkan. Proses ini dilakukan sampai kalimat terakhir, sehingga kata – kata pada kalimat di dokumen bersih dari kata penghubung atau kata berulang – ulang sesuai dengan kamus *stopword* [12]. Kata pada tiap kalimat yang merupakan data hasil *tokenizing* kata sebelum proses *stopword removal* dan sesudah proses *stopword removal* dapat dilihat pada tabel 3.7 sebagai berikut.

Tabel 3. 7 Hasil *Stopword Removal*

	Sebelum						Sesudah			
No	Kata						Kata			
S1	menlo	google	android	dapat	pada		menlo	google	mempercepat	mobile
	park	telah	di	mempercepat	perangkat		park	diluncurkan	proses	
	fitur	diluncurkan	seluruh	proses	mobile		fitur	pengguna	checkout	
	pay	untuk	dunia	checkout			pay	android	online	
	with	pengguna	agar	online			with	dunia	perangkat	
S2	biasanya	beberapa	situs				pengguna	langganan		
	para	formulir	langganan				mengisi	berbayar		
	pengguna	untuk	berbayar				formulir			
	harus	masuk					masuk			
	mengisi	berbagai					situs			
S3	dilansir	pengguna	kartu	dapat	pengguna	chrome	dilansir	1	google	chrome
	Dari	dapat	kredit	ditambahkan	melalui	atau	dailymail	kartu	pengguna	android
	dailymail	memiliki	atau	ke	google	android	pengguna	kredit	google	pay
	sekarang	salah	debit	akun	play	pay	memiliki	debit	play	
	para	1	yang	google	youtube	chrome	salah	akun	youtube	
S4	google	pembayaran	layanan	mereka	mengetikkannya		google	pengiriman	mengetikkannya	
	nantinya	dan	dengan	yang			mengirimkan	pengguna		
	akan	informasi	informasi	berarti			informasi	layanan		
	mengirimkan	pengiriman	dari	tidak			pembayaran	informasi		
	informasi	pengguna	akun	perlu			informasi	akun		
S5	saat	dapat	dengan	sebagai			diluncurkan	sekira		
	diluncurkan	digunakan	sekira	coming			pay	selusin		
	pay	untuk	selusin	soon			with	terdaftar		
	with	15	Lebih				google	coming		
	google	tempat	terdaftar				15	soon		

Tabel 3. 7 Hasil *Stopword Removal* (Lanjutan)

	Sebelum							Sesudah			
No	Kata							Kata			
S6	layanan	eat24	instacart	spot				layanan	planning	service	
	yang	perusahaan	dan	hero				tersedia	kayak	spot	
	tersedia	travel	parking					eat24	instacart	hero	
	sekarang	planning	reservation					perusahaan	parking		
	termasuk	kayak	service					travel	reservation		
S7	mereka	soon	stubhub	deliveroo				mereka	perusahaan	just	gomez
	yang	termasuk	papa	dan				terdaftar	tiket	eat	
	terdaftar	airbnb	johns	guzman				coming	stubhub	deliveroo	
	sebagai	perusahaan	just	y				soon	papa	guzman	
	coming	tiket	eat	gomez				airbnb	johns	y	
S8	fitur	diam	bulan					fitur	i		
	itu	saat	mei					preview	o		
	diberi	konferensi						diam	mei		
	preview	i						diam			
	diam	o						konferensi			
S9	mulai	dapat	banyak	favorit	cepat	produk	google	mulai	situs	pali	google
	hari	mempercepat	aplikasi	anda	pali	google		mempercepat	web	bhat	menulis
	ini	checkout	dan	dengan	bhat	menulis		checkout	favorit	vp	blog
	sekarang	online	situs	beberapa	vp	di		online	klik	manajemen	google
	anda	di	web	klik	manajemen	blog		aplikasi	cepat	produk	
S10	membayar	menggunakan	secara	global				membayar	payment		
	berbagai	fitur	resmi					tagihan	api		
	tagihan	google	telah					goolge	resmi		
	dengan	payment	diluncurkan					fitur	diluncurkan		
	goolge	api	secara					google	global		

Tabel 3. 7 Hasil *Stopword Removal* (Lanjutan)

No	Sebelum						Sesudah		
	Kata						Kata		
S11	fitur	di	chrome				fitur		
	tersebut	aplikasi	google				aplikasi		
	dapat	android					android		
	digunakan	di					chrome		
S12	pihak	google	ketika				pihak	dipungut	
	mengatakan	tidak	menggunakannya				google	biaya	
	bahwa	dipungut					pay	transaksi	
	pay	biaya					with	menggunakannya	
	with	transaksi					google		
S13	fitur	asisten	memberi	orang	teman	meminta	fitur	membayar	pengguna
	tersebut	google	tahu	orang	pengguna	otentikasi	asisten	orang	otentikasi
	juga	sehingga	asisten	tertentu	dan		google	orang	
	bekerja	pengguna	untuk	dari	kemudian		pengguna	daftar	
	dengan	dapat	membayar	daftar	akan		asisten	teman	

3.2.3 Ekstraksi Fitur

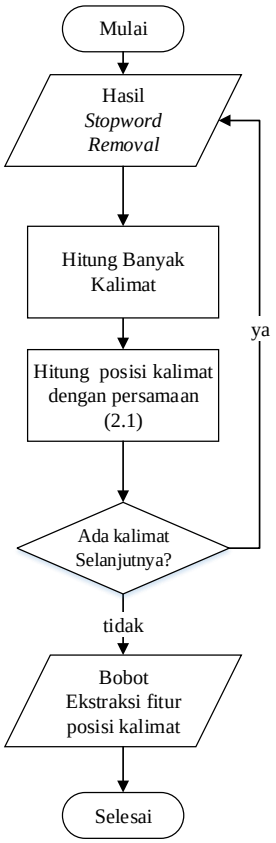
Ekstraksi fitur merupakan proses mengolah kalimat untuk menghasilkan nilai-nilai bobot. Nilai bobot hasil ekstraksi fitur merupakan masukan yang digunakan pada metode pembelajaran *neural network*. Fitur ekstraksi yang digunakan adalah posisi kalimat, *positive keyword*, kemiripan antar kalimat, panjang kalimat dan bobot koneksi antar kalimat. Tahapan dari proses ekstraksi fitur dapat dilihat pada gambar 3.9 sebagai berikut.



Gambar 3. 9 Proses Ekstraksi Fitur

1. Posisi Kalimat

Posisi kalimat adalah letak kalimat dari kalimat pada dokumen. Posisi kalimat diasumsikan bahwa kalimat pertama merupakan kalimat penting. Perhitungan posisi kalimat diperoleh dari persamaan (2.1). Gambaran *flowchart* diagram posisi kalimat dapat dilihat pada gambar 3.10 sebagai berikut.



Gambar 3. 10 *Flowchart* Diagram Ekstraksi Fitur Posisi kalimat

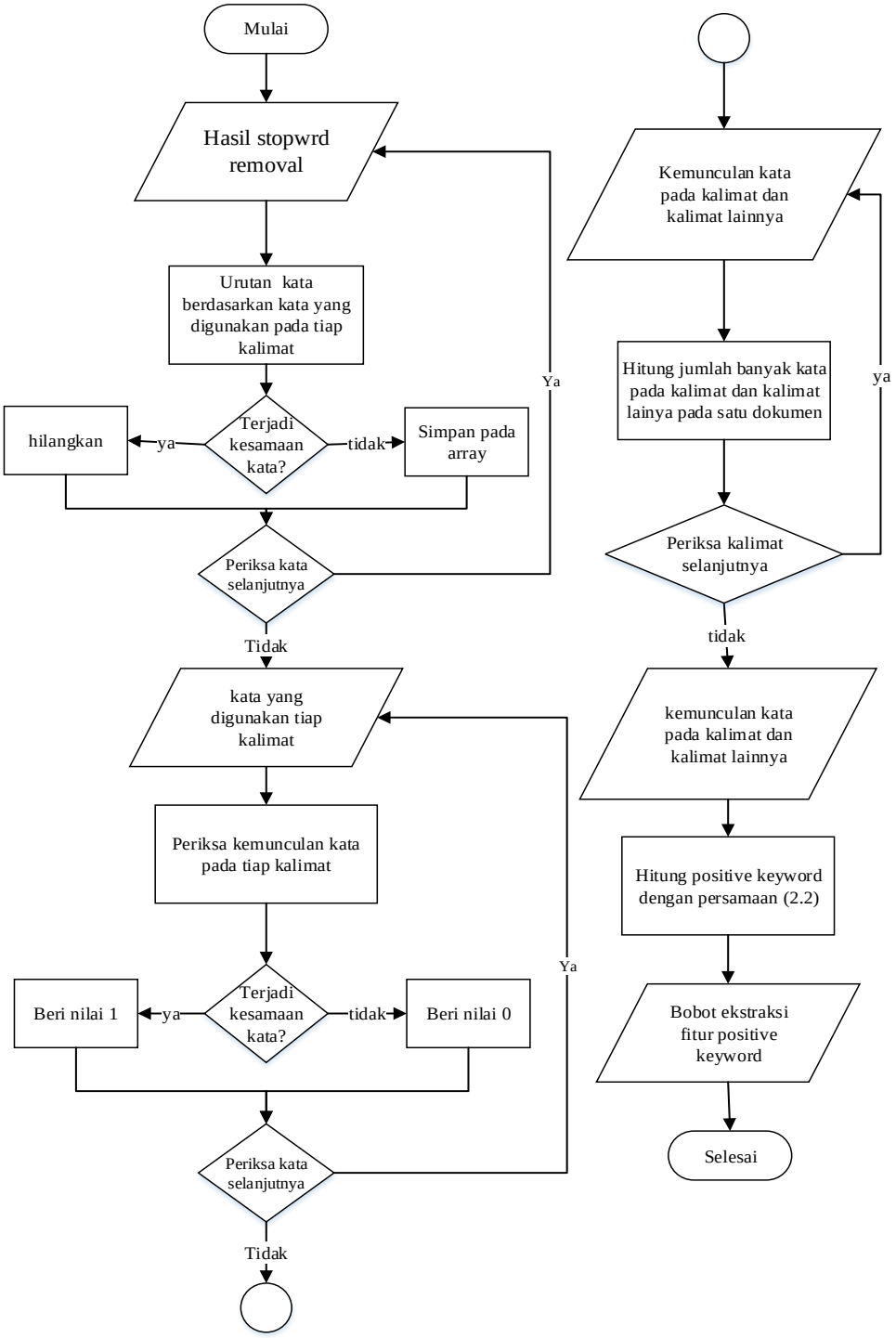
Data hasil *stopword removal* yang terdapat pada tabel 3.7 diproses dengan persamaan (2.1) yaitu perhitungan dari fitur ekstraksi posisi kalimat. Kalimat pertama pada dokumen merupakan nilai yang didapatkan dari banyaknya jumlah pada kalimat dalam dokumen. Kalimat selanjutnya nilai pada kalimat pertama yang merupakan nilai tertinggi dari jumlah kalimat pada satu dokumen, maka kalimat selanjutnya dikurangi 1 dari nilai kalimat sebelumnya yang tertinggi hingga kalimat terakhir bernilai 1. Adapun perhitungan ekstraksi fitur posisi kalimat yang dimana dalam contoh perhitungan kalimat terdiri dari 13 kalimat dalam satu dokumen, maka nilai dari kalimat pertama adalah 13 dan kalimat selanjutnya dikurang 1. Perhitungan dapat dilihat pada tabel 3.8 sebagai berikut.

Tabel 3. 8 Perhitungan Ekstraksi Fitur Posisi Kalimat

$\text{Score}_{f1}(Si)$	$\frac{\text{X}(Nilai\ bobot\ kalimat\ dari\ N)}{\text{N}(Banyaknya\ Kalimat)}$	Bobot Ekstraksi Fitur Posisi kalimat
$\text{Score}_{f1}(s1)$	$\frac{13}{13}$	1
$\text{Score}_{f1}(s2)$	$\frac{12}{13}$	0.923
$\text{Score}_{f1}(s3)$	$\frac{11}{13}$	0.846
$\text{Score}_{f1}(s4)$	$\frac{10}{13}$	0.769
$\text{Score}_{f1}(s5)$	$\frac{9}{13}$	0.692
$\text{Score}_{f1}(s6)$	$\frac{8}{13}$	0.615
$\text{Score}_{f1}(s7)$	$\frac{7}{13}$	0.538
$\text{Score}_{f1}(s8)$	$\frac{6}{13}$	0.462
$\text{Score}_{f1}(s9)$	$\frac{5}{13}$	0.385
$\text{Score}_{f1}(s10)$	$\frac{4}{13}$	0.308
$\text{Score}_{f1}(s11)$	$\frac{3}{13}$	0.231
$\text{Score}_{f1}(s12)$	$\frac{2}{13}$	0.154
$\text{Score}_{f1}(s13)$	$\frac{1}{13}$	0.077

2. *Positive Keyword*

Positive keyword merupakan kata yang sering muncul pada satu kalimat dengan kalimat yang lainnya pada satu dokumen. Banyak kata pada kalimat dan kalimat lainnya pada satu dokumen dapat dilakukan dengan perhitungan *frekuensi* kejadian kata atau kemunculan kata [22]. Gambaran dari *flowchart* diagram *positive keyword* dapat dilihat pada gambar 3.11 sebagai berikut.



Gambar 3. 11 *Flowchart* Diagram *Positive Keyword*

Data hasil *stopword removal* yang terdapat pada tabel 3.7 diproses dengan mengurutkan kata yang digunakan pada tiap kalimat dan selanjutnya melakukan perhitungan *inverse document frequency* untuk mendapatkan nilai kemunculan kata

pada kalimat dan kalimat lainnya dalam satu dokumen. Kata yang muncul pada kalimat dinilai satu meskipun kata dalam kalimat tersebut muncul lebih dari satu kali. Setelah total kalimat dari *positive keyword* didapat, maka tahap selanjutnya perhitungan persamaan *positive keyword* (2.2). Perhitungan dari kemunculan kata pada kalimat dapat dilihat dalam lampiran A-1. Kemunculan kata yang merupakan banyaknya kata yang sering muncul pada tiap kalimat dihitung kemunculan katanya pada tiap kalimat sehingga jumlah kata positive keyword diketahui banyaknya kemunculan pada tiap kalimat, perhitungan dapat dilihat pada lampiran A-1.

Adapun perhitungan *positive keyword* yaitu jumlah kata dalam suatu kalimat mengandung *keyword* dibagi dengan jumlah kata dalam seluruh kalimat yang mengandung *keyword*, dengan *keyword* merupakan banyaknya kata yang muncul dalam suatu dokumen diperoleh dari persamaan (2.2). Berdasarkan perhitungan kemunculan kata yang telah dilakukan pada lampiran A-1 ditemukan bahwa terdapat 20 kata sebagai positive keyword dikarenakan muncul lebih dari 1 kali pada satu dokumen. Kalimat pertama mendapatkan 10 kali kemunculan kata yang sama berdasarkan 20 kata positive keyword, maka perhitungannya adalah 10 dibagi 20 dan bobotnya adalah 0.500. Perhitungan lebih lengkap dari positive keyword dapat dihitung dengan perhitungan yang dapat dilihat pada Tabel 3.9 sebagai berikut.

Tabel 3. 9 Perhitungan Ekstraksi Fitur *Positive Keyword*

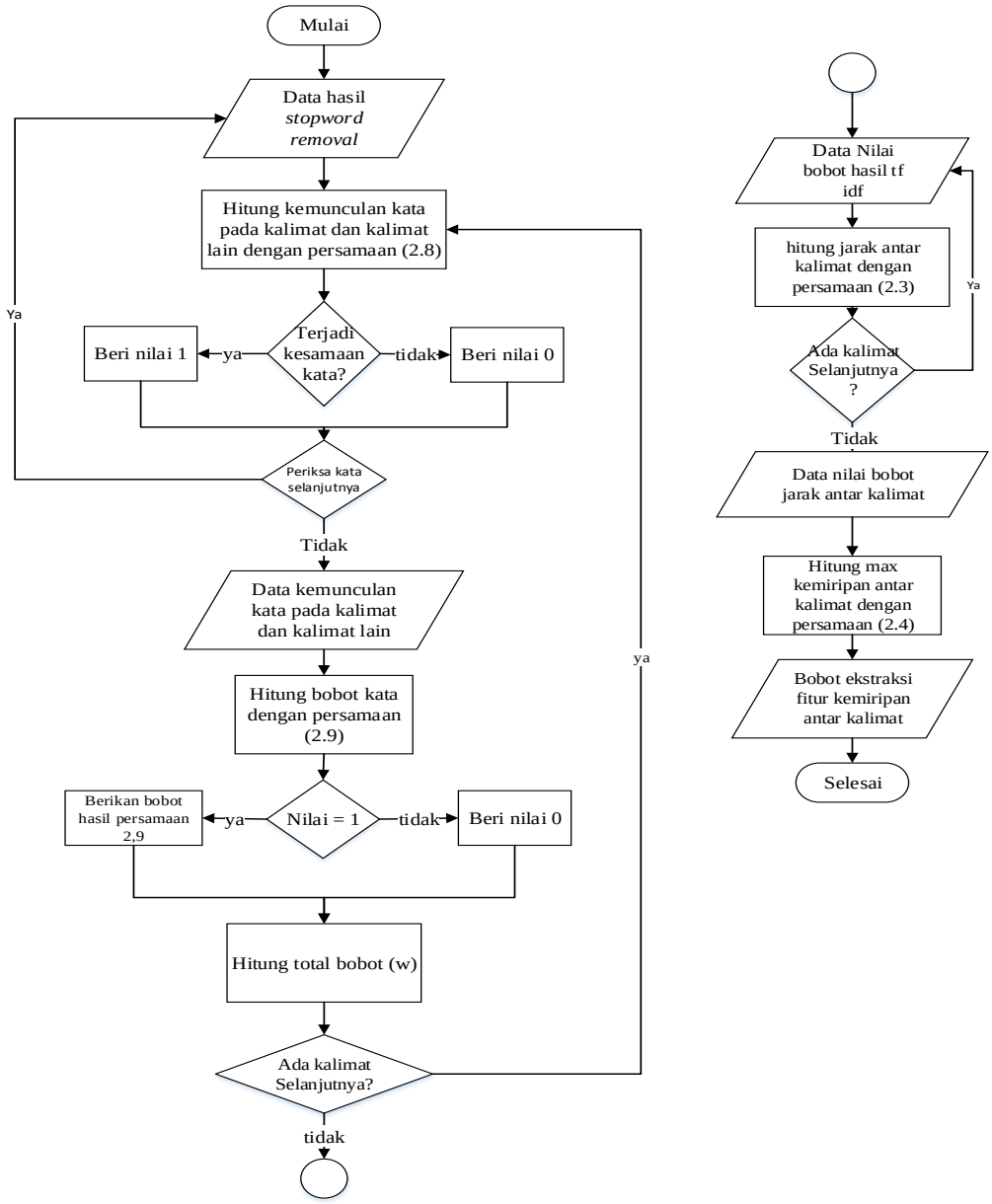
$\text{Score}_{f2}\,(si)$	$\frac{S_i(positif\ keyword)}{\sum_{i=0}^{n} S_i\,(positif\ keword)}$	Bobot Ekstraksi Fitur Positive keyword
$\text{Score}_{f2}\,(s1)$	$\frac{10}{20}$	0.500
$\text{Score}_{f2}\,(s2)$	$\frac{2}{20}$	0.100
$\text{Score}_{f2}\,(s3)$	$\frac{6}{20}$	0.300
$\text{Score}_{f2}\,(s4)$	$\frac{4}{20}$	0.200
$\text{Score}_{f2}\,(s5)$	$\frac{7}{20}$	0.350
$\text{Score}_{f2}\,(s6)$	$\frac{2}{20}$	0.100
$\text{Score}_{f2}\,(s7)$	$\frac{4}{20}$	0.200

Tabel 3. 9 Perhitungan Ekstraksi Fitur *Positive Keyword* (Lanjutan)

$\text{Score}_{f2}\ (si)$	$\frac{S_i(positif\ keyword)}{\sum_{i=0}^{n} S_i\ (positif\ keword)}$	Bobot Ekstraksi Fitur Positive keyword
$\text{Score}_{f2}\ (s8)$	$\frac{1}{20}$	0.050
$\text{Score}_{f2}\ (s9)$	$\frac{6}{20}$	0.300
$\text{Score}_{f2}\ (s10)$	$\frac{4}{20}$	0.200
$\text{Score}_{f2}\ (s11)$	$\frac{4}{20}$	0.200
$\text{Score}_{f2}\ (s12)$	$\frac{2}{20}$	0.100
$\text{Score}_{f2}\ (s13)$	$\frac{4}{20}$	0.200

3. Kemiripan Antar Kalimat

Kemiripan antar kalimat merupakan kata yang muncul dalam kalimat sama dengan kata yang muncul dalam kalimat lain. Nilai dari hasil *filtering* dihitung bobotnya menggunakan perumusan TF-IDF (*term frequency - inverse document frequency)* dengan persamaan 2.8 dan 2.9. Bobot dari hasil perhitungan TF-IDF akan menjadi bobot perhitungan kemiripan antar kalimat dengan menggunakan persamaan 2.3 dan 2.4. Gambaran *flowchart* diagram kemiripan antar kalimat dapat dilihat pada gambar 3.12 sebagai berikut.



Gambar 3. 12 *Flowchart* Diagram Kemiripan Antar Kalimat

Data hasil *stopword removal* yang terdapat pada tabel 3.7 dihitung total kemunculan katan (tf) pada kalimat dan kalimat lainnya pada satu dokumen. Data hasil kemunculan kata dihitung total kemunculan kata (df) untuk mengetahui bobot yang dimiliki kata dengan persamaan (2.8). Bobot kata dihitung IDF untuk mendapatkan total bobot kemunculan kata pada kalimat dalam dokumen yang dimana bobot kalimat (w) tersebut dijadikan data masukan sebagai nilai dalam perhitungan selanjutnya jarak antar kalimat dengan persamaan (2.3). Hasil perhitungan jarak antar kalimat pada kalimat dan kalimat lainnya bobot atau nilai

yang di hasilkan dijumlahkan dan data tersebut menjadi data nilai bobot kalimat. Nilai bobot kalimat yang dihasilkan menjadi masukan untuk perhitungan max kemiripan antar kalimat dengan persamaan (2.4). Proses perhitungan tf-idf pada tiap kalimat yang dimana S1 sampai S13 adalah jumlah kalimat dari kalimat 1 sampai kalimat 13 dapat dilihat pada lampiran A-1.

Nilai bobot dari kemunculan kata pada kalimat dengan kalimat lainnya dalam satu dokumen selanjutnya dihitung kemiripan antar kalimat yang diperoleh dari perhitungan jarak antar kalimat. Perhitungan jarak kalimat diperoleh dari persamaan (2.3). Proses perhitungan jarak antar kalimat dari kalimat 1 sampai kalimat 13 dapat dilihat pada Lampirkan 5.

Bobot dari nilai perhitungan jarak antar kalimat sebelumnya pada lampiran A-1 selanjutnya adalah menghitung kemiripan antar kalimat yaitu dengan bobot kalimat dibagi dengan bobot tertinggi pada kalimat dengan persamaan (2.4). Berdasarkan dari jumlah bobot yang dihasilkan dari perhitungan jarak antar kalimat yang merupakan bobot dari perhitungan TF-IDF pada kalimat dari tiap kalimat dalam satu dokumen pada lampiran A-1 untuk kalimat pertama yaitu sebesar 1.999 dan untuk bobot jarak antar kalimat terbesar pada kalimat sebesar 2.701, maka bobot pada kalimat pertama adalah 1.999 dibagi 2.701 sama dengan 0.740.Hasil perhitungan lebih lengkapnya kemiripan kalimat dapat dilihat pada tabel 3.10 sebagai berikut.

Tabel 3. 10 Perhitungan Ekstraksi Fitur Kemiripan Antar Kalimat

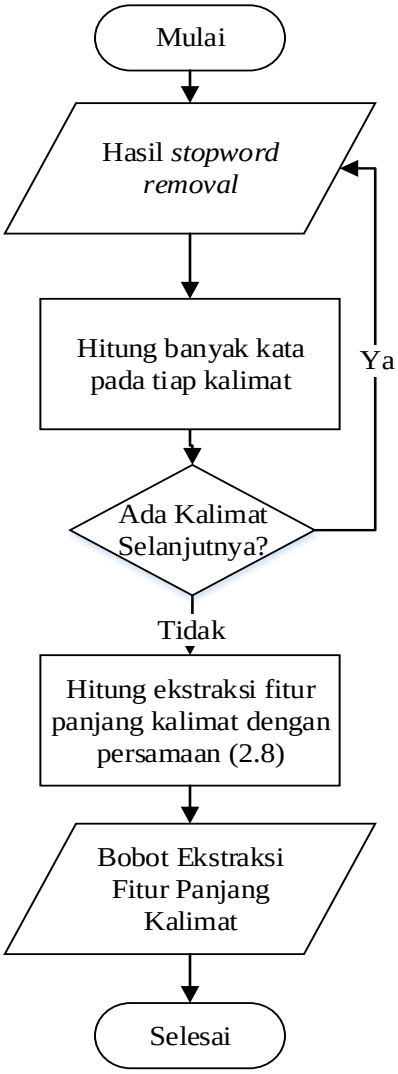
$\text{Score}_{f3}\,(si)$	$\frac{\sum \text{Sim}\,(s_i, s_j)}{\max(\sum \text{Sim}\,(s_i, s_j))}$	Bobot Ekstraksi Fitur Kemiripan Antar Kalimat
$\text{Score}_{f3}\,(S_1)$	$\frac{1.999}{2.701}$	0.74
$\text{Score}_{f3}\,(S_2)$	$\frac{1.364}{2.701}$	0.505
$\text{Score}_{f3}\,(S_3)$	$\frac{1.706}{2.701}$	0.632
$\text{Score}_{f3}\,(S_4)$	$\frac{1.473}{2.701}$	0.545

Tabel 3. 10 Perhitungan Ekstraksi Fitur Kemiripan Antar Kalimat (lanjutan)

$\text{Score}_{f3}\ (si)$	$\frac{\sum \text{Sim}\ (s_i, s_j)}{\max(\sum \text{Sim}\ (s_i, s_j))}$	Bobot Ekstraksi Fitur Kemiripan Antar Kalimat
$\text{Score}_{f3}\ (S_5)$	$\frac{2.031}{2.701}$	0.752
$\text{Score}_{f3}\ (S_6)$	$\frac{1.118}{2.701}$	0.414
$\text{Score}_{f3}\ (S_7)$	$\frac{1.209}{2.701}$	0.448
$\text{Score}_{f3}\ (S_8)$	$\frac{1.999}{2.701}$	0.444
$\text{Score}_{f3}\ (S_9)$	$\frac{1.419}{2.701}$	0.525
$\text{Score}_{f3}\ (S_{10})$	$\frac{1.852}{2.701}$	0.686
$\text{Score}_{f3}\ (S_{11})$	$\frac{2.701}{2.701}$	1
$\text{Score}_{f3}\ (S_{12})$	$\frac{1.957}{2.701}$	0.725
$\text{Score}_{f3}\ (S_{13})$	$\frac{1.720}{2.701}$	0.637

4. Panjang Kalimat

Panjang kalimat merupakan untuk menghilangkan kalimat – kalimat yang terlalu pendek [2]. Dihitung berdasarkan jumlah kata dalam kalimat dibagi jumlah kata terpanjang dalam kalimat di satu dokumen. *Flowchart* diagram dari panjang kalimat dapat dilihat pada gambar 3.12 sebagai berikut.



Gambar 3. 13 *Flowchart* Diagram Panjang Kalimat

Data hasil *stopword removal* yang terdapat pada tabel 3.7 diperoleh sekumpulan kata pada kalimat yang dihitung panjang atau banyak kata pada kalimat satu dan kalimat lainnya dalam satu dokumen. Kalimat dengan kata paling banyak dari kalimat satu dokumen dibagi dengan banyak kata pada suatu kalimat sesuai dengan persamaan (2.8) yang merupakan perhitungan ekstraksi fitur panjang kalimat. Jumlah kata pada masing-masing kalimat diketahui berdasarkan hasil data *stopword* removal yang dapat dilihat pada tabel 3.11 sebagai berikut.

Tabel 3. 11 Panjang Kalimat

kata berdasarkan urutan kalimat												
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13
menlo	pengguna	dilansir	google	diluncurkan	layanan	mereka	fitur	mulai	membayar	fitur	pihak	fitur
park	mengisi	dailymail	mengirimkan	pay	tersedia	terdaftar	preview	mempercepat	tagihan	aplikasi	google	asisten
fitur	formulir	pengguna	informasi	with	eat24	coming	diam	checkout	google	android	pay	google
pay	masuk	memiliki	pembayaran	google	perusahaan	soon	diam	online	fitur	chrome	with	pengguna
with	situs	salah	informasi	15	travel	airbnb	konferensi	aplikasi	google		google	asisten
google	langganan	1	pengiriman	sekira	planning	perusahaan	i	situs	payment		dipungut	membayar
diluncurkan	berbayar	kartu	pengguna	selusin	kayak	tiket	o	web	api		biaya	orang
pengguna		kredit	layanan	terdaftar	instacart	stubhub	mei	favorit	resmi		transaksi	orang
android		debit	informasi	coming	parking	papa		klik	diluncurkan		menggunakannya	daftar
dunia		akun	akun	soon	reservation	johns		cepat	global			teman
mempercepat		google	mengetikkannya		service	just		pali				pengguna
proses		pengguna			spot	eat		bhat				otentikasi
checkout		google			hero	deliveroo		vp				
online		play				guzman		manajemen				
perangkat		youtube				y		produk				
mobile		chrome				gomez		google				
		android						menulis				
		pay						blog				
								google				
16	7	18	11	10	13	16	8	19	10	4	9	12

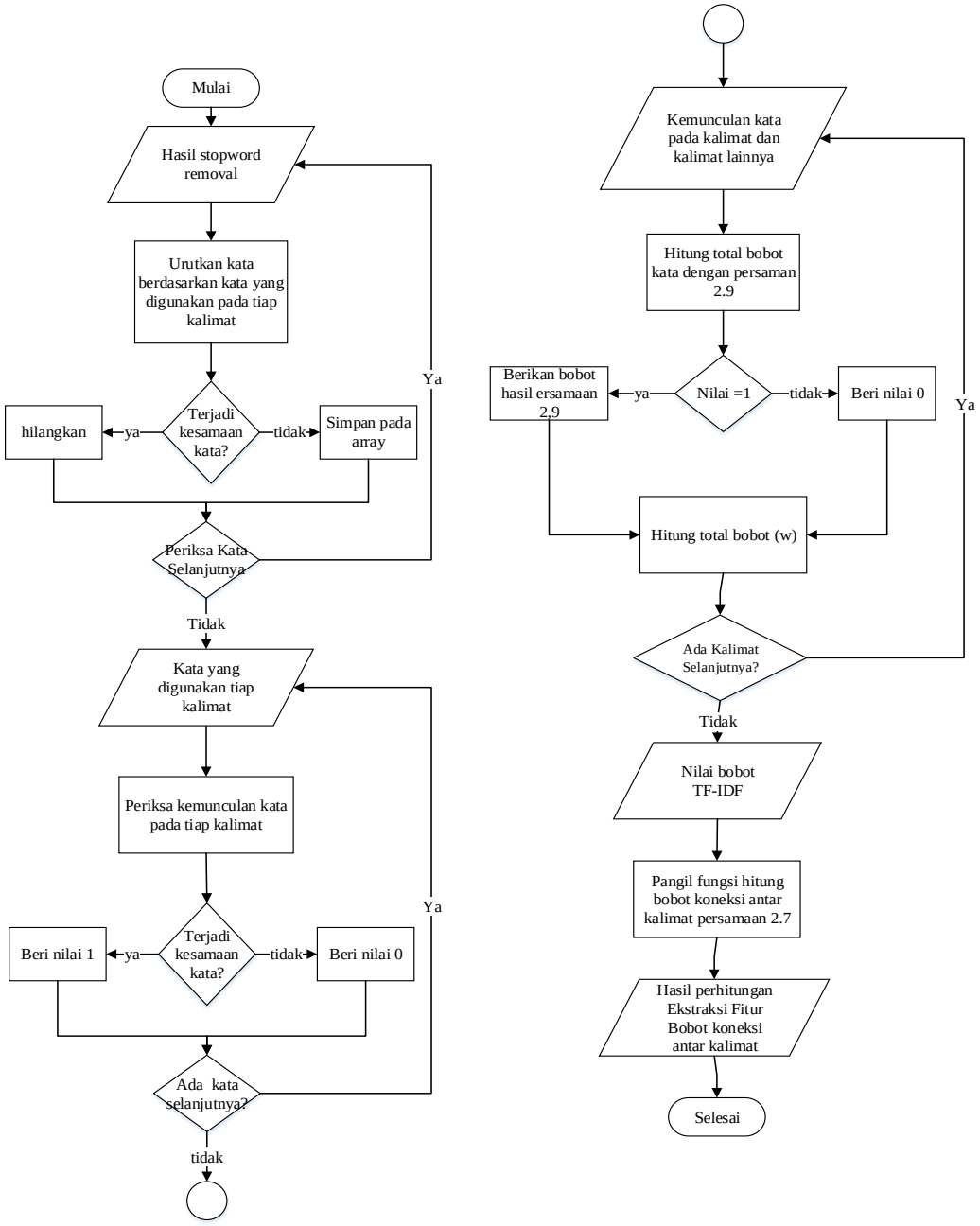
Panjang kalimat didapat dari jumlah kata pada kalimat dibagi dengan jumlah kalimat terpanjang [2]. Tabel diatas menunjukkan nilai kata pada kalimat {16, 7, 18, 11, 10, 13, 15, 8, 18, 10, 4, 8, 12} dan kalimat terpanjang adalah D9 dengan 18 kata. Adapun perhitungan dari panjang kalimat dari persamaan (2.8) dapat dilihat pada tabel 3.12 sebagai berikut.

Tabel 3. 12 Perhitungan Ekstraksi Fitur Panjang Kalimat

Perhitungan Panjang Kalimat		
$\text{Score}_{f4}\,(si)$	$\frac{jumlah\ kata}{Kalimat\ terpanjang}$	Bobot Ekstraksi Fitur Panjang Kalimat
$\text{Score}_{f4}\,(s1)$	$\frac{16}{18}$	0.889
$\text{Score}_{f4}\,(s2)$	$\frac{7}{18}$	0.389
$\text{Score}_{f4}\,(s3)$	$\frac{18}{18}$	1.000
$\text{Score}_{f4}\,(s4)$	$\frac{11}{18}$	0.611
$\text{Score}_{f4}\,(s5)$	$\frac{10}{18}$	0.556
$\text{Score}_{f4}\,(s6)$	$\frac{13}{18}$	0.722
$\text{Score}_{f4}\,(s7)$	$\frac{16}{18}$	0.889
$\text{Score}_{f4}\,(s8)$	$\frac{8}{18}$	0.444
$\text{Score}_{f4}\,(s9)$	$\frac{19}{18}$	1.000
$\text{Score}_{f4}\,(s10)$	$\frac{10}{18}$	0.556
$\text{Score}_{f4}\,(s11)$	$\frac{4}{18}$	0.222
$\text{Score}_{f4}\,(s12)$	$\frac{8}{18}$	0.500
$\text{Score}_{f4}\,(s13)$	$\frac{12}{18}$	0.667

5. Bobot Koneksi Antar Kalimat

Bobot koneksi antar kalimat adalah menjumlahkan bobot kata dalam suatu kalimat dengan kata yang sama dari kalimat lain dalam suatu dokumen [13]. Penjumlahan bobot koneksi antar kalimat digunakan dengan menghitung bobot setiap kalimat dan menghitung semua bobot antar kalimat. *Flowchart* diagram bobot koneksi antar kalimat dapat dilihat pada gambar 3.14 sebagai berikut.



Gambar 3. 14 *Flowchart* Diagram Bobot Koneksi Antar Kalimat

Data hasil *stopword* removal yang terdapat pada tabel 3.7 diurutkan kata – kata yang terdapat pada kalimat dalam dokumen. Kata yang memiliki kemunculan kata yang sama pada kalimat dan kalimat lainnya pada dokumen dihitung satu. Bobot koneksi antar kalimat memerlukan bobot hasil perhitungan berdasarkan TF-IDF yang dimana proses menjumlahkan kata yang digunakan dalam suatu kalimat dengan kata yang sama dari kalimat lain dalam suatu dokumen. Perhitungan kemunculan kata pada kalimat dapat dilihat pada lampiran A-1.

Perhitungan kemunculan kata pada kalimat telah dilakukan. Maka, telah mendapatkan total kemunculan kata dalam kalimat dari dokumen. Pada tahap selanjutnya hitung proses perhitungan bobot kata pada kalimat dari perhitungan kemunculan kata pada kalimat (tf). Proses perhitungan bobot kata pada kalimat dapat dilihat pada lampiran A-1.

Kemunculan dan bobot kata pada kalimat yang dilakukan menghasilkan bobot kalimat yang menjumlahkan bobot kata dalam suatu kalimat dengan kata yang sama dari kalimat lain dalam suatu dokumen dilanjutkan dengan perhitungan persamaan (2.9) untuk mendapatkan bobot koneksi antar kalimat. Berdasarkan perhitungan bobot kata yang digunakan pada satu dokumen dari lampiran A-1 di kalimat pertama mendapatkan jumlah bobot sebesar 12.55 dan bobot tertinggi kalimat dalam satu dokumen adalah 16.47, maka bobot koneksi antar kalimat pada kalimat pertama adalah 12.55 dibagi 16.77 sama dengan 0.762. Adapun perhitungannya lebih lengkapnya dapat dilihat pada Tabel 3.13 sebagai berikut.

Tabel 3. 13 Perhitungan Bobot Koneksi Antar Kalimat

$\text{Score}_{fi}\ (si)$	$\frac{jumlah\ bobot}{bobot\ terbesar}$	Hasil
$\text{Score}_{f5}\ (s1)$	$\frac{12.55}{16.47}$	0.762
$\text{Score}_{f5}\ (s2)$	$\frac{6.77}{16.47}$	0.411
$\text{Score}_{f5}\ (s3)$	$\frac{14.49}{16.47}$	0.88
$\text{Score}_{f5}\ (s4)$	$\frac{7.79}{16.47}$	0.473
$\text{Score}_{f5}\ (s5)$	$\frac{7.76}{16.47}$	0.471

Tabel 3. 13 Perhitungan Bobot Koneksi Antar Kalimat (Lanjutan)

$\text{Score}_{fi}\ (si)$	$\frac{jumlah\ bobot}{bobot\ terbesar}$	Hasil
$\text{Score}_{f5}\ (s6)$	$\frac{13.83}{16.47}$	0.84
$\text{Score}_{f5}\ (s7)$	$\frac{15.45}{16.47}$	0.938
$\text{Score}_{f5}\ (s5)$	$\frac{7.07}{16.47}$	0.429
$\text{Score}_{f5}\ (s9)$	$\frac{16.47}{16.47}$	1
$\text{Score}_{f5}\ (s10)$	$\frac{7.63}{16.47}$	0.463
$\text{Score}_{f5}\ (s11)$	$\frac{2.67}{16.47}$	0.162
$\text{Score}_{f5}\ (s12)$	$\frac{6.80}{16.47}$	0.352
$\text{Score}_{f5}\ (s13)$	$\frac{7.40}{16.47}$	0.449

3.2.4 Nilai Bobot Ekstraksi Fitur

Berdasarkan perhitungan ekstraksi fitur pada kalimat yang telah dilakukan sebelumnya, hasil dari perhitungannya dapat dilihat pada Tabel 3.14.

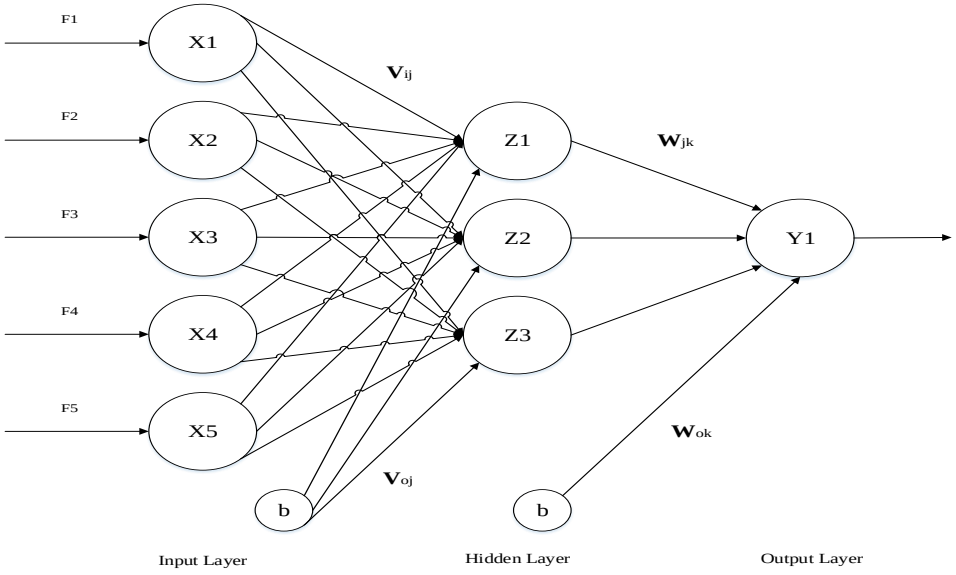
Tabel 3. 14 Hasil Perhitungan Ekstraksi Fitur

Kalimat	Posisi Kalimat	Positive Keyword	Kemiripan Antar Kalimat	Panjang Kalimat	Bobot Koneksi Antar Kalimat
	F1	F2	F3	F4	F5
S1	1	0.500	0.74	0.889	0.762
S2	0.923	0.100	0.505	0.389	0.411
S3	0.846	0.300	0.632	1.000	0.880
S4	0.769	0.200	0.545	0.611	0.473
S5	0.692	0.350	0.752	0.556	0.471
S6	0.615	0.100	0.414	0.722	0.840
S7	0.538	0.200	0.448	0.889	0.938
S8	0.462	0.050	0.444	0.444	0.429
S9	0.385	0.300	0.525	1.000	1.000
S10	0.308	0.200	0.686	0.556	0.463
S11	0.231	0.200	1	0.222	0.162
S12	0.154	0.100	0.725	0.500	0.351
S13	0.077	0.200	0.637	0.667	0.449

3.2.5 Analisis Metode

Algorima *neural network backpropagation* melakukan dua tahap perhitungan yaitu perhitungan maju untuk menghitung *error* antara keluaran *actual* dan target, dan perhitungan mundur yang mempropagasikan baik *error* tersebut untuk memperbaiki bobot *sinapsis* pada semua *neuron* yang ada. Pengujian *backpropagation* dilakukan setelah proses pelatihan *backpropagation* selesai. Fungsi aktivasi yang mentransformasikan dari total input pada suatu neuron untuk menghasilkan sinyal keluaran. Fungsi aktivasi yang digunakan adalah fungsi aktivasi biner.

Arsitektur *backpropagation* terdiri dari tiga lapis yaitu lapisan masukan (*input layer*), lapisan tersembunyi (*hidden layer)* dan lapisan keluaran (*output*). Data yang digunakan sebagai data masukan pada lapisan masukan adalah nilai hasil dari proses ekstraksi fitur (posisi kalimat, *positive keyword*, kemiripan antar kalimat, panjang kalimat dan bobot koneksi antar kalimat) yang diproses tiap kalimat satu persatu hingga kalimat terakhir. Berdasarkan salah satu aturan menentukan *hidden layer* menurut Heaton [18] pada penelitian ini dengan lima input *layer* dan satu *output* jumlah *neuron* pada *hidden layer* adalah tiga *neuron*. Gambaran dari arsitektur *backpropagation* yang digunakan pada gambar 3.15 sebagai berikut.



Gambar 3. 15 Arsitektur *Backpropagation* (maju)

Keterangan:

X = Masukan (menggunakan lima input yang berasal dari nilai hasillima ekstraksi fitur kalimat pada tiap kalimat)

V_{ij} = Nilai bobot pada lapisan tersembunyi X ke unit Z

V_{oj} = Bias pada lapisan tersembunyi X ke unit Z

Z = Bobot pada lapisan tersembunyi (menggunakan tiga neuron)

W_{jk} = Nilai bobot sambungan dari Z ke unit Y

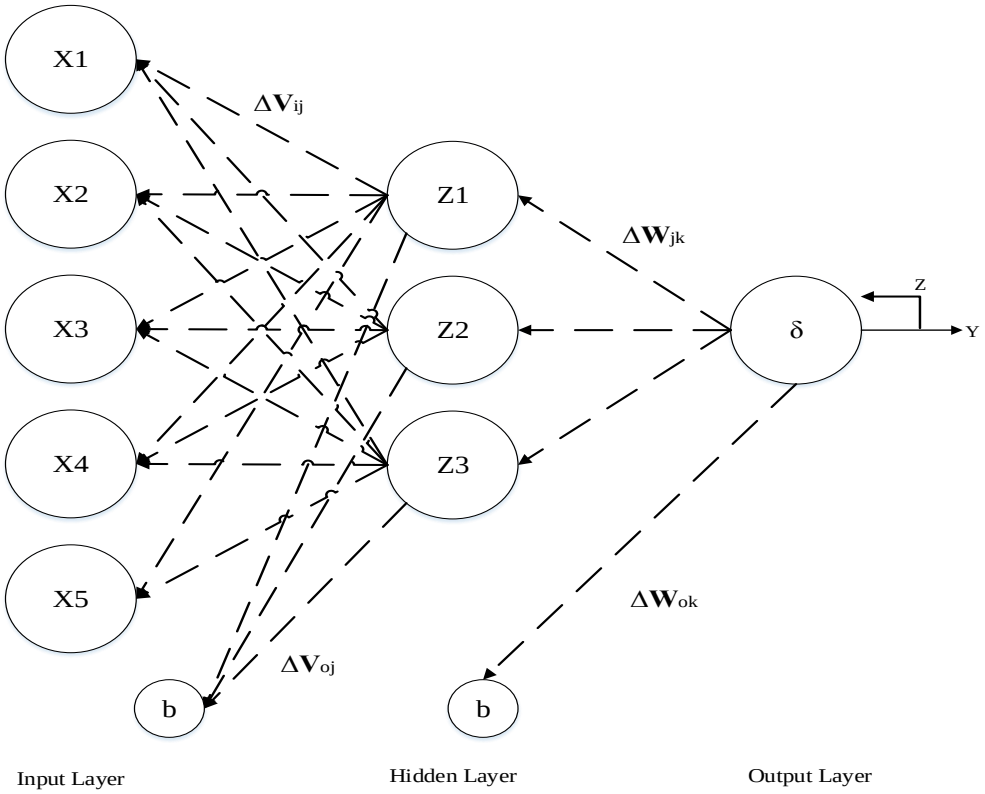
W_{ok} = Bias pada sambungan dari Z ke unit Y

b = Bias pada lapisan tersembunyi dan lapisan keluaran

Y = Keluaran (menggunakan satu output)

F = Fitur ekstraksi

Berdasarkan keterangan di atas adalah arsitektur *backpropagation* untuk propagasi maju dan pada tahap selanjutnya adalah propagasi mundur yang dimana terjadi perbaikan bobot error W_{ij}, W_{0j}, W_{jk} dan W_{0k}. Arsitektur *backpropagation* untuk propagasi mundur pada gambar 3.16 sebagai berikut.



Gambar 3. 16 Arsitektur *Backpropagation* (mundur)

Keterangan:

X_i = Masukan (menggunakan lima input yang berasal dari nilai hasil lima ekstraksi fitur kalimat pada tiap kalimat)

ΔV_{ij} = Nilai bobot pada lapisan tersembunyi X ke unit Z

ΔV_{oj} = Nilai bobot pada lapisan tersembunyi X ke unit Z

Z_i = Bobot pada lapisan tersembunyi (menggunakan tiga neuron)

ΔW_{jk} = Nilai bobot sambungan dari Z ke unit Y

ΔW_{ok} = Bias pada sambungan dari Z ke unit Y

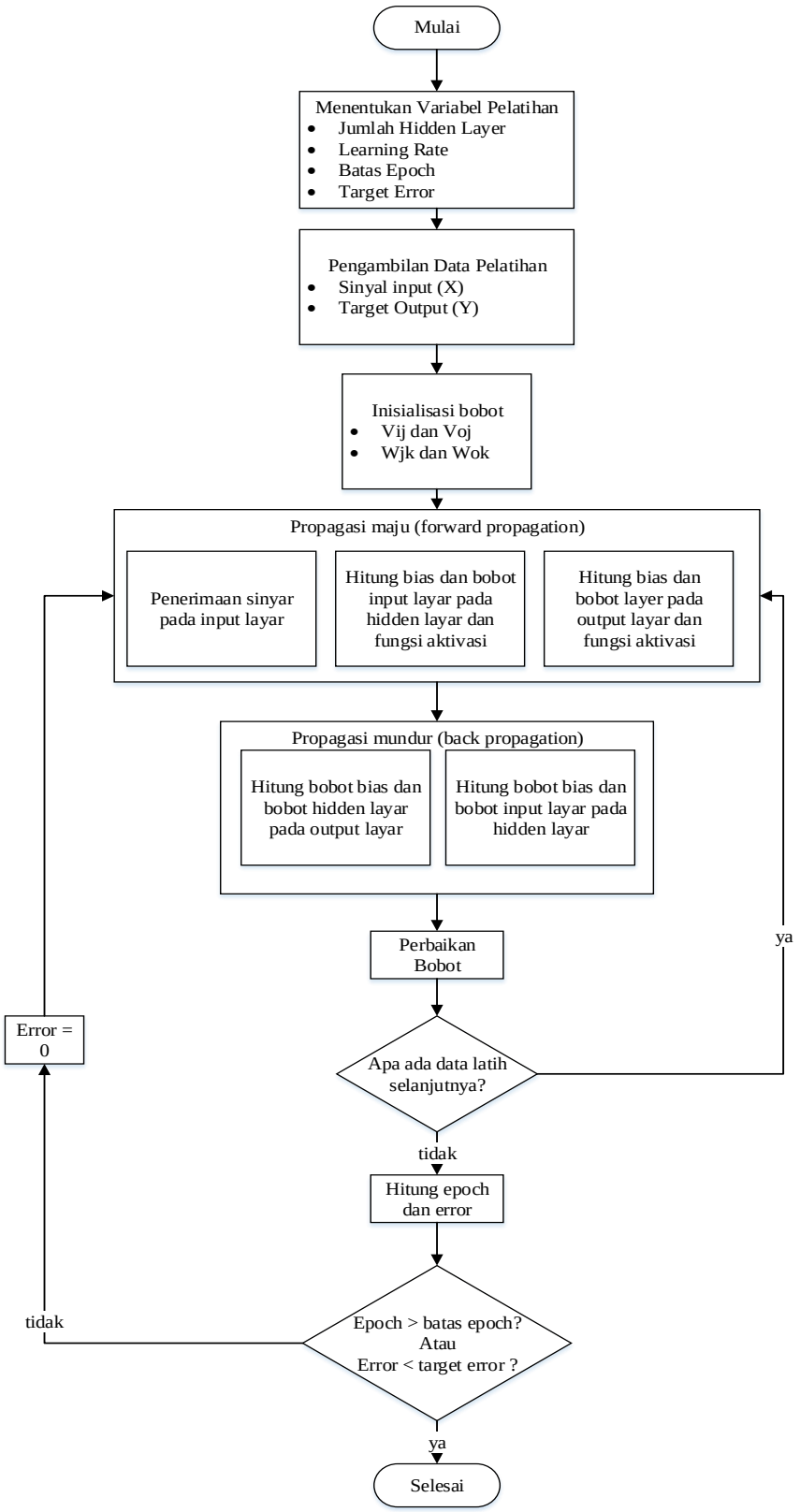
b = Bias pada lapisan tersembunyi dan lapisan keluaran

δ = Faktor pengaturan nilai penimbang sambungan pada lapisan keluaran.

Setelah melakukan proses propagasi mundur bobot lama dan bobot *error* dijumlahkan sehingga menghasilkan bobot baru yang akan digunakan dalam proses selanjutnya adalah kembali ke propagasi maju yang dimana telah menggunakan bobot baru dalam perhitungannya. Proses propagasi maju telah dilakukan maka hitung nilai epoch dan error apakah sudah sesuai dengan yang ditentukan jika belum tercapai maka proses kembali ke proses propagasi mundur untuk diperbaiki kembali dan dihitung kembali.

3.2.5.1 Pelatihan *Backpropagation*

Tahapan pelatihan *backpropagation* dilakukan agar metode dapat mempelajari berbagai pola yang terjadi pada sebuah kasus tersebut. *Block* diagram menggambarkan setiap bagian dari alur kerja. Berikut merupakan tahap – tahap dari proses pelatihan metode pelatihan *backpropagation* yang diimplementasikan pada peringkasan teks otomatis dapat dilihat pada gambar 3.17 sebagai berikut.



Gambar 3. 17 ***Block Diagram Backpropagation***

Tahapan proses berdasarkan *block* diagram *backpropagation* digambar 3.17 diterapkan dalam peringkasan teks otomatis dengan artikel Bahasa Indonesia adalah sebagai berikut.

1. Menentukan variabel pelatihan

Tahap pertama ini, terdapat empat variabel pelatihan yang ditentukan, yaitu jumlah *neuron* pada *hidden layer*, *learning rate*(α), batas epoch, dan batas iterasi (*error*). Penentuan nilai variabel pelatihan dipengaruhi pemilihan jumlah lapisan tersembunyi (*hidden layer*) dan nilai *learning rate* (laju pembelajaran) yang sesuai. Apabila nilai *learning rate* terlalu besar, jaringan akan mencapai *konvergen* dalam waktu yang singkat tetapi *error* klasifikasi menjadi besar. Sebaliknya bila *learning rate* terlalu rendah, hasil klasifikasi memiliki akurasi yang baik tetapi proses *training* membutuhkan waktu yang lama. Batas *epoch* dan target *error* berfungsi untuk membatasi pelatihan, dimana kedua variabel ini dibandingkan dengan kondisi pelatihan. Pada penelitian ini menggunakan variabel pelatihan berdasarkan penelitian Naufal Rachmatullah dan Anggina Primanita [4]. Berikut ini adalah variabel pelatihan yang digunakan pada analisis metode sebagai berikut.

Jumlah neuron pada hidden layer = 3 (berdasarkan arsitektur *backpropagation*)
Learning Rate (α) = 0.5
Batas *Epoch* = 100
Target batas iterasi (*Error*) = 0.01

2. Pengambilan data pelatihan

Data pelatihan yang digunakan pada metode ini berupa nilai bobot hasil dari proses ekstraksi fitur yang telah dilakukan sebelumnya. Kalimat yang telah melalui proses ekstraksi fitur digunakan sebagai input atau sinyal masukan pada *neural network* tiap per-kalimat dari kalimat pertama sampai kalimat terakhir. Terdapat lima sinyal input (X) dan satu *output* (Y). Nilai target output adalah nilai yang berdasarkan dari data latih manual berupa nilai 0 yang diasumsikan sebagai bukan kalimat penting dan nilai 1 yang diasumsikan sebagai kalimat penting. Nilai target *output* digunakan sebagai perhitungan *error* dan perhitungan pada tahapan propagasi mundur. Sinyal input berdasarkan bobot ekstraksi fitur dan target *output* berdasarkan data latih berdasarkan proses manual pada tabel 3.15 sebagai berikut.

Tabel 3. 15 Sinyal Input Hasil Perhitungan Ekstraksi Fitur dan Target Output

Kalimat	F1	F2	F3	F4	F5	Target Output
S1	1	0.500	0.74	0.889	0.762	1
S2	0.923	0.100	0.505	0.389	0.411	0
S3	0.846	0.300	0.632	1.000	0.880	1
S4	0.769	0.200	0.545	0.611	0.473	0
S5	0.692	0.350	0.752	0.556	0.471	1
S6	0.615	0.100	0.414	0.722	0.840	1
S7	0.538	0.200	0.448	0.889	0.938	1
S8	0.462	0.050	0.444	0.444	0.429	0
S9	0.385	0.300	0.525	1.000	1.000	1
S10	0.308	0.200	0.686	0.556	0.463	1
S11	0.231	0.200	1	0.222	0.162	0
S12	0.154	0.100	0.725	0.500	0.351	0
S13	0.077	0.200	0.637	0.667	0.449	0

3. Inisialisasi bobot

Inisialisasi bobot dilakukan untuk seluruh jaringan bobot yang menghubungkan antara *input layer ke*pada *hidden layer* dan *hidden layer ke*pada *output* layer. Bobot yang menghubungkan antara *input layer ke*pada *hidden layer* dengan notasi (V_{ij}) dan biasnya dengan notasi (V_{oj}). Bobot yang menghubungkan antara *hidden layer* kepada *output* layer dilambangkan dengan (W_{jk}) sedangkan untuk biasnya dilambangkan dengan(W_{ok}). Ada dua cara untuk menginisialisasi bobot, yaitu inisialisasi secara random dan inisialisasi menggunakan Nguyen – Widrow. Inisialisasi yang dilakukan pada penelitian ini dengan memberi nilai acak (*random*) terkecil antara 0 hingga 1 ke seluruh bobot dan bias. Contoh dari nilai acak (*random*) yang digunakan pada satu kalimat sebagai contoh dari perhitungan yang akan dilakukan pada analisis metode ini.

$V_{1,1}$= 0.3314 $V_{1,2}$= 0.8349 $V_{1,3}$= 0.8945

$V_{2,1}$= 0.1468 $V_{2,2}$= 0.1411 $V_{2,3}$= 0.1916

$V_{3,1}$= 0.2474 $V_{3,2}$= 0.2413 $V_{3,3}$= 0.6853

$V_{4,1}$= 0.6284 $V_{4,2}$= 0.2615 $V_{4,3}$= 0.1212

$V_{5,1}$= 0.5760 $V_{5,2}$= 0.8032 $V_{5,3}$= 0.7470

$V_{0,k}$= 0.7035

$W_{1,1}$= 0.1110 $W_{1,2}$= 0.4885 $W_{1,3}$= 0.3835

$W_{0,k}$= 0.5700

4. Propagasi maju

Propagasi maju pada tahap ini dilakukan penjumlahan bobot, bias dan fungsi aktivasi pada hasil perjumlahannya. Propagasi maju terdapat tiga tahapan yaitu penerimaan sinyal pada *input layer*, perhitungan bobot *input layer* kepada *hidden layer* dan perhitungan bobot *hidden layer* kepada *output layer.*

a. Penerimaan sinyal pada *input layer*

Penerimaan sinyal pada *input layer* adalah tahapan untuk memasukan nilai data latih kepada setiap neuron pada *input layer*. Contoh penerimaan sinyal pada *input layer* kepada *input layer* pada satu kalimat yang berdasarkan data hasil ekstraksi fitur yang dilakukan sebelumnya dapat dilihat pada tabel 3.16 sebagai berikut.

Tabel 3. 16 **Sinyal Penerimaan Pada Input Layer**

Kalimat	F1	F2	F3	F4	F5	Target Output
S1	1	0.500	0.74	0.889	0.762	1

Data diatas adalah salah satu contoh kalimat yang dimana tiap kalimat di proses berdasarkan nilai bobot ekstraksi fiturnya masing – masing.

b. Hitung bias dan bobot *input layer* pada *hidden layer* dan fungsi aktivasi

Pada langkah ini digunakan nilai *input* ($X_i = X1, .., n$), nilai bobot(V_{ij}) , dan nilai bias(V_{oj}) yang menghubungkan *input layer* kepada *hidden layer* (Z_{inj}), untuk menghitung jumlah bobot pada *input layer* yang menuju kepada *hidden layer* menggunakan persamaan (2.14). Perhitungan *input layer* ke *hidden layer* dari analisis metode sebagai berikut.

Z_{in1} = 0. 7035+ 1*0. 3314+ 0. 500*0. 1468+ 0. 74*0. 2474+ 0. 889*0. 6284+ 0. 762*0. 5760= 2.288

Z_{in2} = 0. 7035+ 1*0. 8349+ 0. 500*0. 1411+ 0. 74*0. 2413+ 0. 889*0. 2615+ 0. 762*0. 8032= 2.632

Z_{in3} = 0. 7035+ 1*0. 8945+ 0. 500*0. 1916+ 0. 74*0. 6853+ 0. 889*0. 1212+ 0. 762*0. 7470= 2.877

Perhitungan tersebut menghasilkan jumlah bobot *input layer* kepada *hidden layer* yang dilambangkan dengan $Z_{inj} = Z_1, Z_2, Z_3$ (merupakan 3 hidden neuron). Nilai Z_{inj} kemudian diaktivasi menggunakan fungsi aktivasi sigmoid biner yang digunakan dengan persamaan (2.11). Perhitungan fungsi aktivasi bobot *input layer* ke *hidden layer* sebagai berikut.

$Z_1 = \frac{1}{1+ exp^{-2.288}}$ = 0.9079

$Z_2 = \frac{1}{1+ exp^{-2.632}}$ = 0.9328

$Z_3 = \frac{1}{1+ exp^{-2.877}}$ = 0.9467

Dari perhitungan Z_{inj} ini dihasilkan nilai dari fungsi aktivasi yang dilambangkan dengan $Z_p = Z_{inj}$ (merupakan bobot sebelum dihitung fungsi aktivitasnya). Nilai $Z_p = Z_1, .., Z_p$ ini digunakan sebagai pengganti sinyal *input* pada langkah selanjutnya yaitu perhitungan *hidden layer* kepada *output layer*.

c. Hitung bias dan bobot *layer* pada *output layer* dan fungsi aktivasi

Perhitungan bias dan bobot pada langkah ini yang menjadi inputan merupakan bobot Z_p yang menghubungkan *hidden layer* kepada *output* dengan menggunakan persamaan (2.16). Perhitungan *input layer* ke *hidden layer* dari analisis metode sebagai berikut.

Dikarenakan jaringan hanya memiliki sebuah jaringan output Y (satu), maka

$$Y_{in1} = 0.5700 + (0.9079 * 0.1110) + (0.9328 * 0.4885) + (0.9467 * 0.3835) = 1.4895$$

Perhitungan tersebut menghasilkan jumlah bobot *hidden layer* kepada *output layer* yang dilambangkan dengan $Y_{ink} = Y_{in1}$ (merupakan 1 output yang digunakan). Nilai Y_{in1} ini kemudian diaktivasi menggunakan fungsi aktivasi sigmoid biner yang digunakan dengan persamaan (2.11). Perhitungan fungsi aktivasi bobot dan bias *input layer* ke *hidden layer* sebagai berikut.

$Y_1 = \frac{1}{1+ exp^{-1.4895}}$ = 0.8160

Perhitungan dari $Y_1 = Y_k$ (merupakan nilai output Y sebelum dihitung dengan fungsi aktifitasnya) yang dihasilkan merupakan nilai yang dugunakan untuk menghitung *error* pada perhitungan mengoreksi bobot pada tahapan propagasi mundur.

5. Propagasi mundur

Propagasi mundur adalah tahapan perhitungan *error* yang digunakan untuk mengoreksi seluruh bobot awal. Pada tahap propagasi mundur melakukan perhitungan bias dan bobot *hidden layer* kepada *output layer* dan perhitungan bias dan bobot *input layer kepada hidden layer*.

a. Hitung *error* bobot bias dan bobot *hidden layer* pada *output layer*

Pada langkah ini digunakan nilai output hidden layer ($Y_k = Y$), nilai Y_k merupakan hasil perhitungan dari bias dan bobot *hidden layer* kepada *output layer* serta fungsi aktivasinya. Menghitung error bobot antara *hidden layer* kepada *output layer* menggunakan persamaan (2.18). Perhitungan *error* bobot *hidden layer* kepada *output layer* sebagai berikut.

$$\delta_1 = (1 - 0.8160) * 0.8160 * (1 - 0.8160) = 0.0276$$

Dari perhitungan tersebut didapatkan nilai *error* pada bobot yang menghubungkan *hidden layer* dengan *output layer* yang dilambangkan dengan $\delta_k = \delta_1$ (karena hanya satu output). Nilai δ_k yang didapatkan proses perhitungan selanjutnya adalah perhitungan koreksi bobot dan bias yang menghubungkan *hidden layer* kepada *output layer* dengan menggunakan persamaan (2.19) dan (2.20). Pada perhitungan koreksi bobot dan bias ini digunakan nilai *error* (δ_k), nilai *learning rate*($\alpha = 0.5$) dan nilai hasil perhitungan fungsi aktivasi yang merupakan penjumlahan dari bias dan bobot *input layer* kepada *hidden layer* ($Z_p = Z_1, Z_2, Z_3$). Perhitungan koreksi bobot dan bias yang menghubungkan *hidden layer* kepada *output layer* sebagai berikut.

Menghitung koreksi bobot *hidden layer*

$$\Delta w_{1,1} = 0.5 * 0.0276 * 0.8943 = 0.0123$$

$$\Delta w_{1,2} = 0.5 * 0.0276 * 0.9220 = 0.0127$$

$$\Delta w_{1,3} = 0.5 * 0.0276 * 0.9382 = 0.0129$$

Menghitung koreksi bias

$$\Delta w_{1,0} = 0.5 * 0.0276 * 0.5700 = 0.0078$$

Berdasarkan perhitungan yang dilakukan, nilai hasil perhitungan tersebut yang dimana bobot *hidden layer* kepada output layer yang dilambangkan $\Delta w_{jk} = \Delta w_{1,1}, \Delta w_{1,2}, \Delta w_{1,3}$ dan bias dengan dilambangkan $\Delta w_{ok} = \Delta w_{0,1}$. Nilai tersebut digunakan untuk memperbaiki bias dan bobot yang menghubungkan hidden layer kepada output layer menjadi bobot baru pada iterasi selanjutnya.

b. Hitung *error* bobot bias dan bobot *input layer* pada *hidden layer*

Setelah mendapatkan nilai koreksi bias dan bobot dari hidden layer kepada *output layer* selanjutnya adalah menghitung koreksi bias dan bobot dari *input layer* pada *hidden layer*. Perhitungan ini dibutuhkan nilai *error* pada bobot penghubung *input layer* kepada *hidden layer* dengan menggunakan persamaan (2.21). Perhitungan *error* bobot *input layer* kepada *hidden layer* sebagai berikut.

$$\delta_{in_1} = 0.0123 * 0.1110 = 0.0013$$

$$\delta_{in_2} = 0.0123 * 0.4885 = 0.0060$$

$$\delta_{in_3} = 0.0123 * 0.3835 = 0.0047$$

Nilai error bobot *input layer* kepada *hidden layer* diketahui dari hasil perhitungan sebelumnya kemudian dikalikan dengan turunan fungsi aktivitasnya untuk mengetahui *error* dari bobot *input layer* kepada hidden layer menggunakan persamaan (2.22). Perhitungan informasi error bobot *input layer* kepada *hidden layers* sebagai berikut.

$$\delta_1 = 0.0013 * 0.8943 * (1 - 0.8943) = 0.00012$$

$$\delta_1 = 0.0060 * 0.9220 * (1 - 0.9220) = 0.00043$$

$$\delta_1 = 0.0047 * 0.9382 * (1 - 0.9382) = 0.00027$$

Setelah mendapatkan informasi *error* bobot *input layer* kepada *hidden layer*, maka proses selanjutnya adalah proses perhitungan perbaikan bias dan bobot yang menghubungkan *input layer* dan *input layer*V_{ij}. Pada perhitungan ini digunakan nilai *error* (δ_k), nilai *learning rate* (α) , dan nilai hasil perhitungan fungsi aktivasi yang merupakan penjumlahan dari bias dan bobot *input layer* kepada *hidden layer* $X_i = X_1 \dots n$ dengan menggunakan persamaan (2.23) untuk

perhitungan bobot dan persamaan (2.24) untuk perhitungan bias. Perhitungan koreksi bobot dan bias input layer kepada hidden layer dapat dilihat pada tabel 3.17 sebagai berikut.

Tabel 3. 17 Koreksi Bobot Dan Bias Input Layer Kepada Hidden Layer

	$\Delta V_{i,1}$	$\Delta V_{i,2}$	$\Delta V_{i,3}$
X1	$0.5 * 0.00012 * 1 = 0.0001471$	$0.5 * 0.00043 * 1 = 0.0002202$	$0.5 * 0.00027 * 1 = 0.0001436$
X2	$0.5 * 0.00012 * 0.842 = 0.0000157$	$0.5 * 0.00043 * 0.842 = 0.0000236$	$0.5 * 0.00027 * 0.842 = 0.0000154$
X3	$0.5 * 0.00012 * 0.724 = 0.0001125$	$0.5 * 0.00043 * 0.724 = 0.0001684$	$0.5 * 0.00027 * 0.724 = 0.0001098$
X4	$0.5 * 0.00012 * 0.724 = 0.0001238$	$0.5 * 0.00043 * 0.724 = 0.0001854$	$0.5 * 0.00027 * 0.724 = 0.0001209$
X5	$0.5 * 0.00012 * 0.721 = 0.0000937$	$0.5 * 0.00043 * 0.721 = 0.0001402$	$0.5 * 0.00027 * 0.721 = 0.0000915$
Bias	$0.5 * 0.00012 = 0.0001471$	$0.5 * 0.00043 = 0.0002202$	$0.5 * 0.00027 = 0.0001436$

Perhitungan tersebut didapatkan nilai perbaikan bias ΔV_{oj} dan bobot ΔV_{ij} yang menghubungkan *input layer* kepada *hidden layer*. Nilai dari perhitungan tersebut digunakan untuk memperbaiki bias dan bobot yang menghubungkan input layer kepada *hidden layer* menjadikan bobot baru yang akan digunakan untuk iterasi selanjutnya.

2. Perbaikan bobot

Perbaikan bobot adalah proses yang dilanjutkan setelah proses propagasi mundur. Nilai yang didapatkan dari proses propagasi mundur dijadikan sebagai nilai perbaikan seluruh bobot. Dalam tahapan ini seluruh nilai bias dan bobot lama ($V_{ij}, V_{oj}, W_{jk}\ dan\ W_{ok}$) dan seluruh nilai hasil dari perbaikan bias dan bobot ($\Delta V_{ij}, \Delta_{oj}, \Delta W_{jk}\ dan\ \Delta W_{ok}$) sebagai perhitungan nilai bias dan bobot yang baru. Bias dan bobot baru yang dihasilkan dengan menambahkan bias dan bobot lama dengan bias dan bobot perbaikan dengan persamaan (2.25) dan (2.26). Perhitungan dari perbaikan bias dan bobot *hidden layer* kepada *output layer* dan *input layer* kepada *hidden layer* dapat dilihat pada tabel 3. 18 sebagai berikut.

Tabel 3. 18 Koreksi Bias Dan Bobot *Hidden Layer* Kepada *Output Layer*

	$V_{i,1(baru)}$	$V_{i,2(baru)}$	$V_{i,3(baru)}$
X1	$0.3314 + 0.0001471 = 0.3315$	$0.8349 + 0.0002202 = 0.8351$	$0.8945 + 0.0001436 = 0.8946$
X2	$0.1468 + 0.0000157 = 0.1468$	$0.1411 + 0.0000236 = 0.1411$	$0.1916 + 0.0000154 = 0.1916$
X3	$0.2474 + 0.0001125 = 0.2476$	$0.2413 + 0.0001684 = 0.2414$	$0.6853 + 0.0001098 = 0.6854$
X4	$0.6284 + 0.0001238 = 0.6285$	$0.2615 + 0.0001854 = 0.2617$	$0.1212 + 0.0001209 = 0.1213$
X5	$0.5760 + 0.0000937 = 0.5761$	$0.8032 + 0.0001402 = 0.8034$	$0.7470 + 0.0000915 = 0.7470$
	$W_{j,1(baru)}$	$W_{j,1(baru)}$	$W_{j,1(baru)}$
Z	$0.1110 + 0.0125 = 0.7036$	$0.4885 + 0.0129 = 0.7037$	$0.3835 + 0.0131 = 0.7036$
	$V_{0,1(baru)}$	$V_{0,2(baru)}$	$V_{0,3(baru)}$
Bias	$0.7035 + 0.0001471 = 0.3749$	$0.7035 + 0.0002202 = 0.3274$	$0.7035 + 0.0001436 = 0.5827$
	$W_{01(baru)}$		
Bias	$0.5700 + 0.0080 = 0.5780$		

Seluruh nilai bias dan bobot yang baru dari perhitungan tersebut, maka pelambangan rumus yang sama dengan bias dan bobot sebelumnya yaitu $V_{ij}, V_{oj}, W_{jk}\ dan\ W_{ok}$. Nilai baru dari bias dan bobot digunakan sebagai iterasi selanjutnya.

3. Pemeriksaan data latih

Pemeriksaan data latih merupakan kondisi nilai bobot baru telah didapatkan. Pemeriksaan data latih memiliki kondisi yang harus dipenuhi pada tahapan ini adalah seluruh dari data pelatihan sudah dilatih. Jika tidak kondisi terpenuhi, proses kembali pada tahapan propagasi maju hingga kondisi terpenuhi dan jika kondisi terpenuhi maka, pelatihan akan dilanjutkan pada proses selanjutnya.

4. *Epoch* dan *error*

Epoch dan *error* merupakan tahapan untuk melakukan pemeriksaan kondisi akhir dari *epoch* dan error yang di tentukan. Terdapat dua kondisi yang harus terpenuhi yaitu jika jumlah *epoch* (iterasi) telah melebihi batas yang ditentukan dan kondisi nilai error telah mencapai target *error*. Ada tiga macam perhitungan error

yang sering digunakan, yaitu *Mean Square Error* (MSE), *Mean Absolute Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE).

Mean Square Error (MSE) merupakan perhitungan error yang digunakan dengan rata – rata kuadrat dari selisih dari hasil pengaruh nilai target output T_k dan hasil perhitungan bobot output layer Y_k. Perhitungan dari *error* yang digunakan dalam memperbandingkan kondisi menggunakan persamaan (2.27) sehingga contoh perhitungannya sebagai berikut.

$$MSE = \frac{1}{5}\ (1 -\ 0.4241)^2 = 0.066332162$$

Hasil dari *error* dijumlahkan dan dikuadratkan yang nantinya dibandingkan dengan target *error* (Target batas iterasi = 0.01) dan nilai *learning rate* ($\alpha = 0.5$) yang sudah ditentukan. Pada tahapan ini jika salah satu kondisi terpenuhi maka proses dari pelatihan berhenti dan jika kondisi tidak terpenuhi maka proses pelatihan akan kembali ke tahapan propagasi maju.

3.2.5.2 Pengujian *Backpropagation*

Pengujian *backpropagation*, pada proses pengujian ini menggunakan proses pengujian pada metode *backpropagation* yaitu hanya melalui proses propagasi maju untuk mengidentifikasi pola yang sudah ditentukan. *Block* diagram dalam *backpropagation* dapat dilihat pada gambar 3.18 sebagai berikut.



Gambar 3. 18 *Block* Diagram Pengujian Metode *Backpropagation*

Proses pengujian *backpropagation* yang dimana tahapannya sesuai dengan block diagram pada gambar 3.18 sebagai berikut.

1. Nilai bobot pelatihan *backpropagation*

Pada pengujian *backpropagation*, bobot dan bias yang digunakan adalah bobot dan bias hasil dari pelatihan yang sudah melewati iterasi sampai akhir pelatihan. Contoh setelah *epoch* ke 60 hasil nilai bobot V_{ij}, bias V_{oj}, bobot W_{jk}, dan bias V_{ok} yang diperoleh dari pelatihan sebagai berikut.

$V_{1,1}$= -0.4772 $V_{1,2}$= 0.3289 $V_{1,3}$= -0.0184

$V_{2,1}$= 0.4656 $V_{2,2}$ = -0.3587 $V_{2,3}$= 0.1969

$V_{3,1}$= -0.4548 $V_{3,2}$= -0.3726 $V_{3,3}$= -0.2992

$V_{4,1}$= -0.3938 $V_{4,2}$= 0.3403 $V_{4,3}$= 0.475

$V_{5,1}$= 0.0683 $V_{5,2}$= -0.1912 $V_{5,3}$= 0.5105

$V_{0,1}$= -0.3218 $V_{0,2}$= -0.1269 $V_{0,3}$= 0.2809

$W_{1,1}$= 0.2603 $W_{1,2}$= 0.7343 $W_{1,3}$= 0.5177

$W_{0,1}$= 1.5108

2. Propagasi maju

Propagasi maju pada tahap ini dilakukan penjumlahan bobot, bias dan fungsi aktivasi pada hasil perjumlahannya. Propagasi maju terdapat tiga tahapan yaitu penerimaan sinyal pada *input layer*, perhitungan bobot *input layer* kepada *hidden layer* dan perhitungan bobot *hidden layer* kepada *output layer.*

a. Penerimaan sinyal pada *input layer*

Penerimaan sinyal pada *input layer* adalah tahapan untuk memasukan nilai data latih kepada setiap neuron pada *input layer*. Contoh penerimaan sinyal pada *input layer* kepada *input layer* pada satu kalimat yang berdasarkan data sinyal input pada pelatihan *backpropagation* yang dilakukan sebelumnya dapat dilihat pada tabel 3.19 sebagai berikut.

Tabel 3. 19 Sinyal Penerimaan Pada Input Layer

Kalimat	F1	F2	F3	F4	F5	Target Output
S1	1	0.500	0.74	0.889	0.762	1

b. Hitung bias dan bobot *input layer* pada *hidden layer* dan fungsi aktivasi

Pada langkah ini digunakan nilai *input* (X_1, ... , n), nilai bobot(V_{ij}) , dan nilai bias(V_{0j}) yang menghubungkan *input layer* kepada *hidden layer* (Z_{inj}), untuk menghitung jumlah bobot pada *input layer* yang menuju kepada *hidden layer* menggunakan persamaan (2.14). Perhitungan tersebut menghasilkan jumlah bobot *input layer* kepada *hidden layer* yang dilambangkan dengan Z_{inj}. Nilai Z_{inj} kemudian diaktivasi menggunakan fungsi aktivasi sigmoid biner yang digunakan dengan persamaan (2.11). Perhitungan Z_{inj} ini dihasilkan nilai dari fungsi aktivasi

yang dilambangkan dengan Z_p. Nilai Z_p ini digunakan sebagai pengganti sinyal *input* pada langkah selanjutnya yaitu perhitungan *hidden layer* kepada *output layer*.

Z_{in1} = -0.3218 + 1 * (-0.4772) + 0. 115* 0.4656 + 0. 74* (-0.4548) + 0. 889* (-0.3938) + 0. 762* 0.0683 = -1.3851753

Z_{in2} = -0.1269 + 1 * 0.3289 + 0. 115* (-0.3587) + 0. 74*(-0.3726) + 0. 889* 0.3403 + 0. 762* (-0.1912) = 0.0196095

Z_{in3} = 0.2809 + 1 * (-0.0184) + 0. 115* 0.1969 + 0. 74* (-0.2992) + 0. 889* 0.475 + 0. 762* 0.5105 = 0.8431208

Fungsi aktivasi sigmoid biner (2.11)

$$Z_1 = \frac{1}{1+ exp^{-(-1.3851753)}} = 0.79982089$$

$$Z_2 = \frac{1}{1+ exp^{-0.01960958}} = 0.495097782$$

$$Z_3 = \frac{1}{1+ exp^{-0.8431208}} = 0.300877915$$

c. Hitung bias dan bobot *layer* pada *output layer* dan fungsi aktivasi

Perhitungan bias dan bobot pada langkah ini yang menjadi input merupakan bobot Z_p yang menghubungkan *hidden layer* kepada *output* dengan menggunakan persamaan (2.17). Perhitungan tersebut menghasilkan jumlah bobot *hidden layer* kepada *output layer* yang dilambangkan dengan Y_{ink}yang menghasilkan Y_k, dikarnakan Y_k hanya satu output maka Y_1. Nilai Y_1 ini kemudian diaktivasi menggunakan fungsi aktivasi sigmoid biner yang digunakan dengan persamaan (2.11). Perhitungan dari Y_1 yang dihasilkan merupakan nilai yang dugunakan untuk menjadi kandidat dalam ringkasan.

$$Y_{in1} = 1.5108 + (0.79982089 * 0.2603) + (0.495097782 * 0.7343) + (0.300877915 * 0.5177) = 2.238308176$$

Perhitungan tersebut menghasilkan jumlah bobot *hidden layer* kepada *output layer* yang dilambangkan dengan $Y_{ink} = Y_{in1}$ (merupakan 1 output yang digukan). Nilai Y_{in1} ini kemudian diaktivasi menggunakan fungsi aktivasi sigmoid biner yang digunakan dengan persamaan (2.11). Perhitungan fungsi aktivasi bobot dan bias *input layer* ke *hidden layer* sebagai berikut.

$$Y_1 = \frac{1}{1+ exp^{-2.238308176}} = 0.9864$$

3. Identifikasi Hasil

Setelah memperoleh nilai aktivasi $Y_k = Y_1$ hasil propagasi maju, maka nilai aktivasi akan dibandingkan dengan *threshold* yang bernilai 0.5. Nilai *output* Y_1 lebih besar dari 0.5 maka *output* yang diberikan nilai 1 dan jika nilai *output* lebih kecil dari 0,5 diberikan nilai 0. Berikut ini adalah contoh nilai aktivasi Y_k yang dihasilkan dari perhitungan sebelumnya sebagai berikut.

$$Y_1 = 0.9864$$

Nilai aktivasi Y_k dengan hasil $0.9864 = 1$, dikarenakan $0.9864 > 0.5$. Sesuai dengan target keluaran yang diharapkan yaitu 1. Proses ini akan dilakukan pada kalimat lainnya hingga kondisi terpenuhi.

3.3 Hasil Ringkasan

Pemilihan kalimat sebagai kalimat ringkasan dilakukan jika proses dalam pengujian *neural network* telah selesai dilakukan. Kalimat yang telah didapatkan nilai bobot akhir dijadikan sebagai kandidat kalimat ringkasan dan diurutkan berdasarkan nilai bobot yang tertinggi sampai ke paling rendah. Pemilihan kalimat ringkasan menjadi kalimat ringkasan berdasarkan bobot yang paling tinggi. Kalimat-kalimat hasil ringkasan kemudian disusun kembali berdasarkan urutan aslinya pada teks. Hasil dari perhitungan *neural network* dapat dilihat pada table 3.20 sebagai berikut.

Tabel 3. 20 Hasil Pengujian *Backpropagation*

Kalimat	Output
S9	0.95075115035188
S7	0.93383999590257
S6	0.89267055314276
S3	0.78704764892306
S1	0.57244591875701
S10	0.46213631264078
S5	0.36965609408637
S2	0.36345361126273
S4	0.31420509418147
S13	0.25634656394499
S8	0.1731077799976
S12	0.16289652841925
S11	0.14027143166925

Berdasarkan tabel di atas kalimat dengan nilai output berdasarkan urutan dari nilai tertinggi adalah kalimat S9, S7, S6, S3, S1, S10, S5, S2, S4, S13, S8, S12 dan S11. Hasil ringkasan diambil dari 30% dari kalimat dengan nilai tertinggi maka kalimat yang tertinggi adalah S9, S7, S6 dan S3. Kalimat disusun kembali berdasarkan urutan seperti pada teks aslinya. Hasil dari peringkasan yang disusun kembali berdasarkan kalimat penyusunnya pada tabel 3.21 sebagai berikut.

Tabel 3. 21 Hasil Ringkasan

Artikel Asli	Hasil Ringkasan
MENLO PARK - Fitur Pay with Google telah diluncurkan untuk pengguna Android di seluruh dunia agar dapat mempercepat proses checkout online pada perangkat mobile. Biasanya, para pengguna harus mengisi beberapa formulir untuk masuk berbagai situs langganan berbayar.Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay. Google nantinya akan mengirimkan informasi pembayaran dan informasi pengiriman pengguna layanan dengan informasi dari akun mereka yang berarti tidak perlu mengetikkannya. Saat diluncurkan, Pay with Google dapat digunakan untuk 15 tempat, dengan sekira selusin lebih terdaftar sebagai 'coming soon'. Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero. Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez. Fitur itu diberi preview diam-diam saat konferensi I / O bulan Mei. "Mulai hari ini sekarang Anda dapat	Dilansir dari Dailymail, sekarang para pengguna dapat memiliki salah 1 kartu kredit atau debit yang dapat ditambahkan ke Akun Google pengguna melalui Google Play, YouTube, Chrome, atau Android Pay. Layanan yang tersedia sekarang termasuk Eat24, perusahaan travel planning Kayak, Instacart, dan parking reservation service Spot Hero. Mereka yang terdaftar sebagai 'coming soon' termasuk AirBnB, perusahaan tiket StubHub, Papa Johns, Just Eat, Deliveroo dan Guzman y Gomez. Mulai hari ini sekarang Anda dapat mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat, Pali Bhat, VP manajemen produk Google menulis di blog Google.

mempercepat checkout online di banyak aplikasi dan situs web favorit Anda dengan beberapa klik cepat," Pali Bhat, VP manajemen produk Google menulis di blog Google. Membayar berbagai tagihan dengan Google menggunakan fitur Google Payment API secara resmi telah diluncurkan secara global. Fitur tersebut dapat digunakan di aplikasi Android di Chrome. Pihak Google mengatakan bahwa Pay with Google tidak dipungut biaya transaksi ketika menggunakannya. Fitur tersebut juga bekerja dengan asisten Google sehingga pengguna dapat memberi tahu asisten untuk membayar orang-orang tertentu dari daftar teman pengguna dan kemudian akan meminta otentikasi.	

3.4 Analisis Kebutuhan Sistem

Analisis kebutuhan sistem adalah spesifikasi kebutuhan sistem yang didapatkan berdasarkan kebutuhan perangkat lunak peringkasan teks otomatis dengan metode *neural network* pada artikel berita berbahasa Indonesia. Analisis kebutuhan perangkat lunak terbagi menjadi dua bagian, yaitu kebutuhan analisis kebutuhan non fungsional dan analisis kebutuhan fungsional.

3.4.1 Analisis Kebutuhan Non Fungsional

Analisis kebutuhan non fungsional merupakan sebuah gambaran kebutuhan pendukung perangkat lunak yang diperlukan untuk menjalankan perangkat lunak yang dibangun. Analisis kebutuhan non fungsional terdiri dari analisis perangkat lunak (*software*), analisis kebutuhan perangkat keras (*hardware*) dan analisis kebutuhan pengguna (*user*).

3.4.1.1 Analisis Perangkat Lunak

Terdapat beberapa perangkat lunak yang digunakan dalam pembangunan perangkat lunak serta mengoperasikan perangkat lunak peringkasan teks otomatis dengan metode *neural network* pada artikel berita berbahasa Indonesia. Perangkat

lunak yang digunakan dalam pembangunan peringkasan teks otomatis dapat dilihat pada tabel 3.22 sebagai berikut.

Tabel 3. 22 Analisis Perangkat Lunak Untuk Pembangunan

No	Jenis Perangkat Lunak	Perangkat Lunak
1	Sistem Operasi	Windows 32 bit atau 64 bit
2	*Web Browser*	Mozila Fire Fox, Chrome, IE
3	*Tools*	XAMPP
4	*Text Editor*	*Sublim Text* versi 3

Adapun perangkat lunak yang dapat digunakan bagi penggunaan peringkasan teks otomatis dapat dilihat pada tabel 3.23 sebagai berikut.

Tabel 3. 23 Analisis Perangkat Lunak Untuk Penggunaan

No	Jenis Perangkat Lunak	Perangkat Lunak
1	Sistem Operasi	Windows 32 bit atau 64 bit
2	*Web Browser*	Mozila Fire Fox, Chrome, IE

3.4.1.2 Analisis Kebutuhan Perangkat Keras

Analisis kebutuhan perangkat keras dalam pembangunan serta pengujian dalam mengoperasikan perangkat lunak peringkasan teks otomatis dengan metode *neural network* pada artikel berita berbahasa Indonesia ini dibutuhkan seperangkat komputer. Spesifikasi komputer berdasarkan spesifikasi persyaratan sistem analisis perangkat lunak yang digunakan di tabel 3.27 dan 3.28 dengan memiliki spesifikasi minimum perangkat keras seperti pada tabel 3.24 sebagai berikut.

Tabel 3. 24 Analisis Kebutuhan Perangkat Keras

No	Perangkat Keras	Spesifikasi
1	Processor	Pentium 4
2	Memori	4 GB
3	Harddisk	120 GB
4	DirectX	DirectX 9

3.4.1.3 Analisis Kebutuhan Pengguna

Analisis kebutuhan pengguna merupakan uraian mengenai siapa saja yang dapat menggunakan sistem dan terlibat dalam pengelolaan data. Hal ini, dapat diketahui karakteristik pengalaman dan pemahaman pengguna terhadap perangkat

lunak pada proses peringkasan teks otomatis dengan metode *neural network* pada artikel berita berbahasa Indonesia dapat dilihat pada tabel 3.25 sebagai berikut.

Tabel 3. 25 Analisis Kebutuhan Pengguna

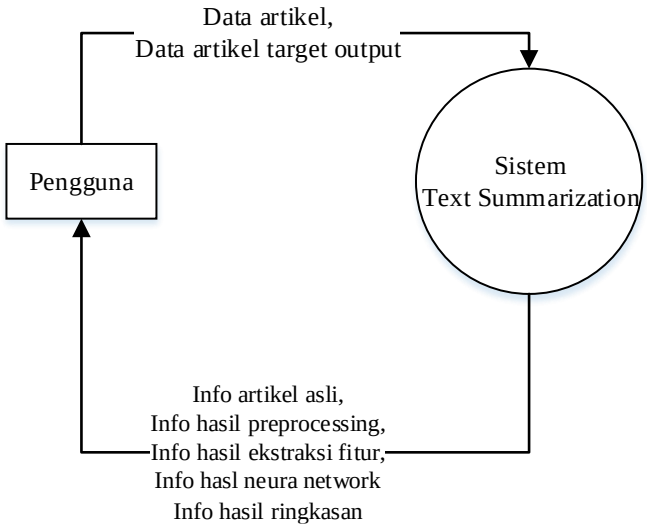
No	Keterangan
1	Menguasai penggunaan komputer
2	Menguasai penggunaan web browser
3	Memahami pengoperasian peringkasan teks

3.4.2 Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional merupakan spesifikasi kebutuhan yang dilakukan perangkat lunak ketika diimplementasikan. Analisis ini diperlukan untuk menentukan masukan yang diperlukan, keluaran yang dihasilkan dan lingkup proses yang digunakan untuk mengelola masukan menjadi keluaran. Adapun analisis kebutuhan fungsional dalam peringkasan teks otomatis dengan metode *neural network* pada artikel berita berbahasa Indonesia adalah sebagai berikut.

3.4.2.1 Diagram Konteks

Diagram konteks untuk Peringkasan teks otomatis pada artikel Bahasa Indonesia dengan menggunakan metode *neural network*, dapat dilihat pada gambar 3.19 sebagai berikut.



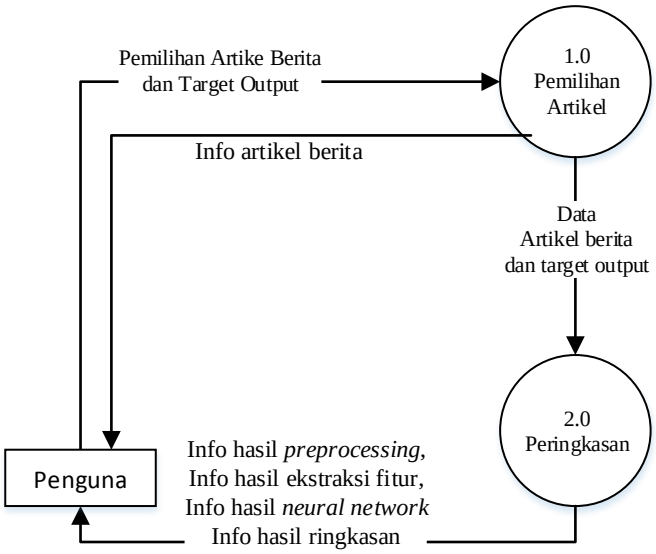
Gambar 3. 19 Diagram Konteks

3.4.2.2 *Data Flow Diagram*

Data flow diagram (DFD) merupakan model sistem yang memperlihatkan fungsional dimana tiap transformasi menunjukkan fungsi atau proses. Berikut ini merupakan *data flow diagram* untuk sistem yang terjadi dalam peringkasan teks otomatis dengan metode *neural network* pada artikel berita bahasa Indonesia.

3.4.2.3 *Data flow diagram* (DFD) Level 1

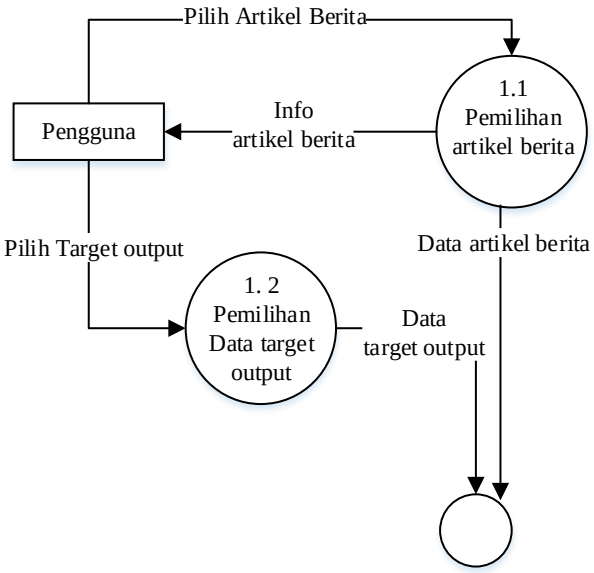
Pada DFD level 1 ini menjelaskan bagian – bagian yang dapat dilakukan pengguna terhadap sistem, dapat dilihat pada gambar 3.20 sebagai berikut.



Gambar 3. 20 DFD Level 1

3.4.2.4 *Data flow diagram* (DFD) Level 2 Pemilihan Artikel

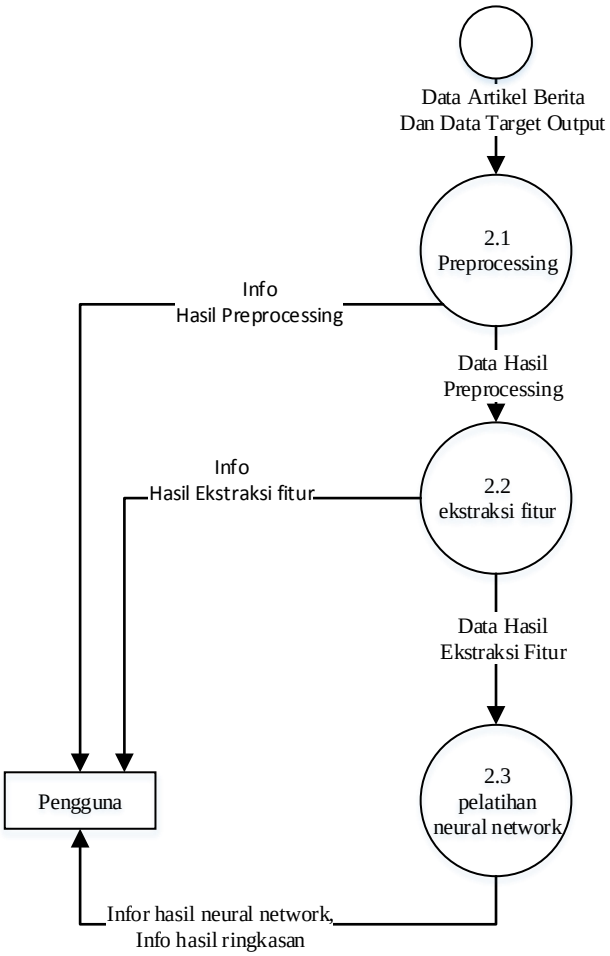
Data flow diagram (DFD) Level 2 menjelaskan aliran data yang terdapat pada proses pemilihan artikel. DFD level 2 pemilihan artikel dapat dilihat pada gambar 3.21 sebagai berikut.



Gambar 3. 21 DFD Level 2 Pemilihan Artikel

3.4.2.5 *Data flow diagram* (DFD) Level 2 Peringkasan

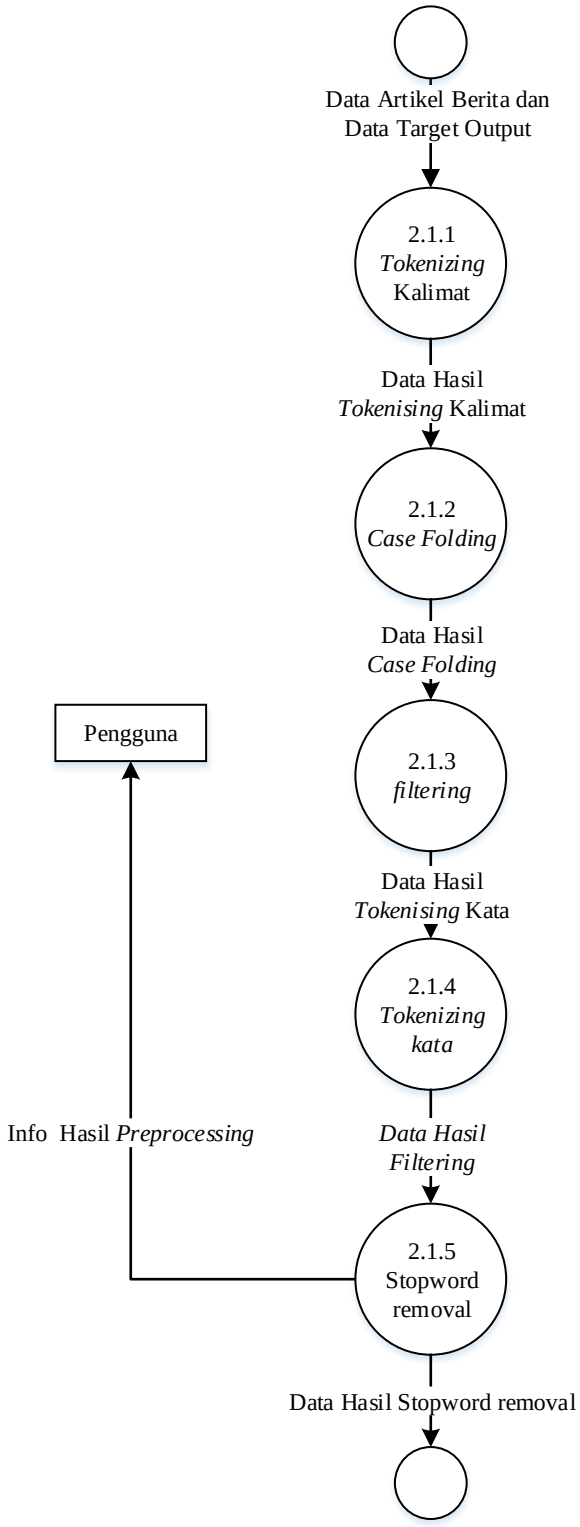
Data flow diagram (DFD) Level 2 menjelaskan aliran data yang terdapat pada proses peringkasan. DFD level 2 peringkasan dapat dilihat pada gambar 3.22 sebagai berikut.



Gambar 3. 22 DFD Level 2 Peringkasan

3.4.2.6 *Data flow diagram* (DFD) Level 3 *preprocessing*

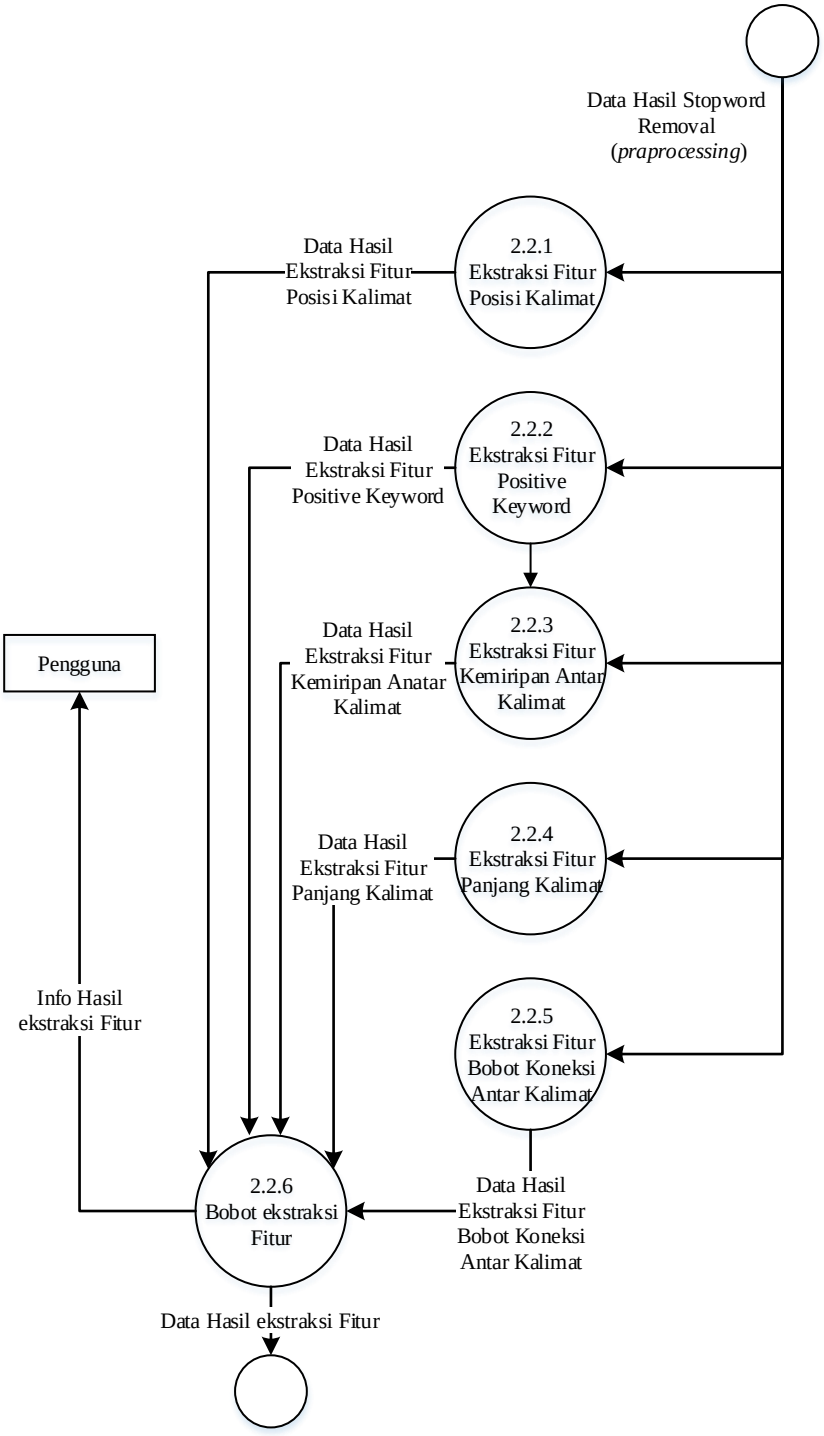
Data flow diagram (DFD) Level 3 menjelaskan aliran data yang terdapat pada proses *preprocessing*. DFD level 3 *preprocessing* dapat dilihat pada gambar 3.23 sebagai berikut.



Gambar 3. 23 DFD Level 3 *Preprocessing*

3.4.2.7 *Data flow diagram* (DFD) Level 3 Ekstraksi Fitur

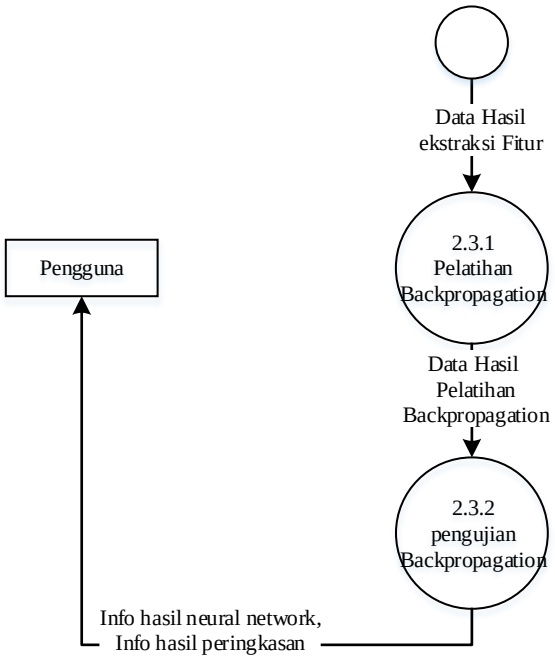
Data flow diagram (DFD) Level 3 menjelaskan aliran data yang terdapat pada proses ekstraksi fitur. DFD level 3 ekstraksi fitur dapat dilihat pada gambar 3.24 sebagai berikut.



Gambar 3. 24 DFD Level 3 Ekstraksi fitur

3.4.2.8 *Data flow diagram* (DFD) Level 3 Pelatihan *Neural Network*

Data flow diagram (DFD) Level 3 menjelaskan aliran data yang terdapat pada proses pelatihan *neural network*. DFD level 3 pelatihan *neural network* dapat dilihat pada gambar 3.25 sebagai berikut.



Gambar 3. 25 DFD Level 3 Pelatihan *Neural Network*

3.4.2.9 Spesifikasi Proses

Spesifikasi proses merupakan keterangan atau deskripsi dari semua proses yang terdapat di DFD. Berikut ini adalah spesifikasi yang merupakan spesifikasi proses untuk DFD pada subbab sebelumnya dapat dilihat pada tabel 3.25 sebagai berikut.

Tabel 3. 26 Spesifikasi Proses

Nomor	Nama Proses	Keterangan
1	Nomor Proses	1.1
	Nama Proses	Pemilihan artikel berita
	Sumber	Pengguna
	Input	Data artikel berita

	Output	Teks artikel berita
	Logika Proses	1. Pemilihan artikel berita 2. Menampilkan data artikel berita yang dipilih berupa teks
2	Nomor Proses	1.2
	Nama Proses	Pemilihan target output
	Sumber	Pengguna
	Input	Data target output
	Output	Target output artikel
	Logika Proses	1. Pemilihan data target output sesuai dengan nama data artikel berita
3	Nomor Proses	2.1.1
	Nama Proses	*Tokenizing* kalimat
	Sumber	Pengguna
	Input	Data teks artikel berita
	Output	Data teks artikel per-kalimat
	Logika Proses	1. *Tokenizing* kalimat berdasarkan *regex*, jika terdapat dalam *regex* kalimat tidak dipisahkan dan jika tidak terdapat dalam *regex* kalimat di-*tokenizing* kalimat berdasarkan *delimiter* tanda titik “.” tanya “ ? ”, dan tanda seru “!”
4	Nomor Proses	2.1.2
	Nama Proses	*Case folding*
	Sumber	Pengguna
	Input	Data hasil *tokenizing* kalimat
	Output	Data teks artikel per-kalimat dengan karakter kecil
	Logika Proses	1. Mengubah karakter kalimat data hasil *tokenizing* kalimat menjadi kalimat dengan karakter huruf kecil
5	Nomor Proses	2.1.3
	Nama Proses	*Filtering*

	Sumber	Pengguna
	Input	Data hasil *case folding*
	Output	Data teks artikel per-kalimat tampa *delimiter*
	Logika Proses	1. Lakukan penghilangan karakter yang dianggap *delimiter* berdasarkan tabel 2.3
6	Nomor Proses	2.1.4
	Nama Proses	*Tokenizing* kata
	Sumber	Pengguna
	Input	Data hasil *filtering*
	Output	Kata per-kalimat
	Logika Proses	1. Lakukan pemotongan kalimat menjadi sekumpulan kata – kata berdasarkan *delimiter* yang menyusunnya, yaitu spasi (" ").
7	Nomor Proses	2.1.5
	Nama Proses	*Stopword removal*
	Sumber	Pengguna
	Input	Data hasil *tokenizing* kata
	Output	Kata per-kalimat
	Logika Proses	1. Kata diperiksa berdasarkan kamus stopwords, jika terdapat dalam kamus kata dihilangkan dan jika tidak terdapat dalam kamus kata tidak dihilangkan
9	Nomor Proses	2.2.1
	Nama Proses	Ekstraksi fitur posisi kalimat
	Sumber	Pengguna
	Input	Data hasil *stopword removal*
	Output	Nilai bobot ekstraksi fitur posisi kalimat per-kalimat
	Logika Proses	1. Hitung banyak kalimat 2. Berikan nilai pada kalimat berdasarkan banyak kalimat yang dimana kalimat berikutnya dikurangi satu (-1)

		3. Hitung bobot ekstraksi fitur posisi kalimat dengan persamaan 2.1 pada tiap kalimat dan kalimat lainnya dalam satu dokumen
10	Nomor Proses	2.2.2
	Nama Proses	Ekstraksi fitur *positive keyword*
	Sumber	Pengguna
	Input	Data hasil *stopword removal*
	Output	Nilai bobot ekstraksi fitur *positive keyword* per-kalimat
	Logika Proses	1. Hitung kemunculan kata pada kalimat dan kalimat lainya pada satu dokumen dengan perhitungan *term frequency* 2. Hitung jumlah kata pada kalimat dan kalimat lainya pada satu dokumen 3. Hitung jumlah kata pada satu dokumen 4. Hitung bobot ekstraksi fitur *positive keyword* dengan persamaan 2.2 pada tiap kalimat dan kalimat lainnya dalam satu dokumen
11	Nomor Proses	2.2.3
	Nama Proses	Ekstraksi fitur kemiripan antar kalimat
	Sumber	Pengguna
	Input	Data hasil *stopword removal*
	Output	Nilai Bobot ekstraksi fitur kemiripan antar kalimat per-kalimat
	Logika Proses	1. Hitung nilai bobot kata pada kalimat dan kalimat lain dengan perhitungan bobot TF – IDF pada persamaan 2.8 dan 2.9 2. Hitung jarak antar kalimat dari nilai bobot TF – IDF dengan persaman 2.4

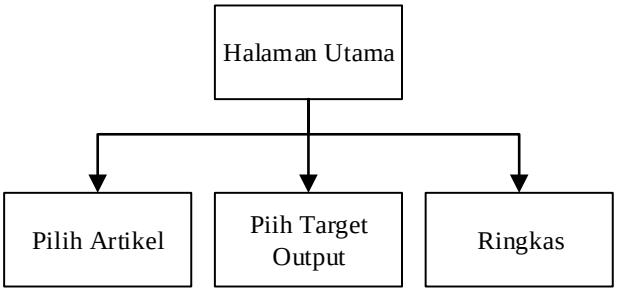
		3. Hitung bobot ekstraksi fitur kemiripan antar kalimat dengan persamaan 2.3 pada tiap kalimat dan kalimat lainnya dalam satu dokumen
12	Nomor Proses	2.2.4
	Nama Proses	Ekstraksi fitur panjang kalimat
	Sumber	Pengguna
	Input	Data hasil *stopword removal*
	Output	Nilai bobot ekstraksi fitur panjang kalimat per-kalimat
	Logika Proses	1. Hitung banyak kata pada tiap kalimat dalam satu dokumen 2. Hitung bobot ekstraksi fitur panjang kalimat dengan persamaan 2.5 pada tiap kalimat dan kalimat lainnya dalam satu dokumen
13	Nomor Proses	2.2.5
	Nama Proses	Ekstraksi fitur bobot koneksi antar kalimat
	Sumber	Pengguna
	Input	Data hasil *stopword removal*
	Output	Nilai bobot ekstraksi fitur bobot koneksi antar kalimat per-kalimat
	Logika Proses	1. Hitung bobot kata yang digunakan pada kalimat dan kalimat lainnya dalam satu dokumen dengan perhitungan bobot TF – IDF dengan persamaan 2.8 dan 2.9 2. Hitung bobot ekstraksi fitur bobot koneksi antar kalimat dengan persamaan 2.6 pada tiap kalimat dan kalimat lainnya dalam satu dokumen
14	Nomor Proses	2.3.1
	Nama Proses	Pelatihan *backpropagation*
	Sumber	Pengguna
	Input	Bobot nilai ekstraksi fitur

	Output	Nilai bobot hasil pelatihan *input layer*, *hidden layer* dan bias
	Logika Proses	1. Ambil data data bobot nilai ekstraksi fitur per-kalimat 2. Pelatihan dengan satu kalimat, jika selesai pelatihan proses dilakukan kembali hingga kalimat terakhir 3. Inisialiasi bobot 4. Hitung bias dan bobot *input layer* pada *hidden layer* dan fungsi aktivasi dengan persamaan 2.14 dan 2.11 5. Hitung bias dan bobot hidden *layer* pada *output layer* dan fungsi aktivasi dengan persamaan 2.16 dan 2.11 6. Hitung *error* bobot bias dan bobot *hidden layer* pada *output layer* dengan persamaan 2.18 7. Hitung koreksi bobot bias dan bobot *hidden layer* pada *output layer* dengan persamaan 2.19 dan 2.20 8. Hitung *error* bobot bias dan bobot *input layer* pada *hidden layer* dengan persamaan 2.21 9. Hitung koreksi bobot bias dan bobot *input layer* pada *hidden layer* dengan persamaan 2.24 dan 2.25 10. Hitung bobot koreksi dari bobot lama dan bobot baru dengan persamaan 2.26 dan 2.27 11. Hitung nilai MSE dengan persamaan 2.27 12. Iterasi akan berhenti jika nilai learning rate terpenuhi
15	Nomor Proses	2.3.2
	Nama Proses	pengujian *backpropagation*

	Sumber	Pengguna
	Input	Nilai bobot hasil pelatihan *backpropagation*
	Output	Nilai bobot hasil pengujian *output* layer per-kalimat
	Logika Proses	1. Ambil data hasil pelatihan *backpropagation* yaitu bobot terakhir hasil iterasi 2. Hitung bias dan bobot *input layer* pada *hidden layer* dan fungsi aktivasi dengan persamaan 2.14 dan 2.11 3. Hitung bias dan bobot *hidden layer* pada *output layer* dan fungsi aktivasi dengan persamaan 2.16 dan 2.11 4. Hitung bobot *output layer* jika sesuai target output maka yang dimana bobot sesuai dengan *learning rate* proses berhenti dan jika tidak sesuai bobot kembali dilatih hingga bobot terpenuhi 5. Urutkan nilai bobot output layer dari nilai tertinggi sampai terendah

3.5 Perancangan Struktur Menu

Perancangan struktur menu merupakan bagian – bagian yang dapat digunakan oleh user pada aplikasi atau program. Berikut ini merupakan perancangan struktur menu yang dibuat dalam pembangunan *prototype* peringkasan teks otomatis dapat dilihat pada gambar 3.26 sebagai berikut.



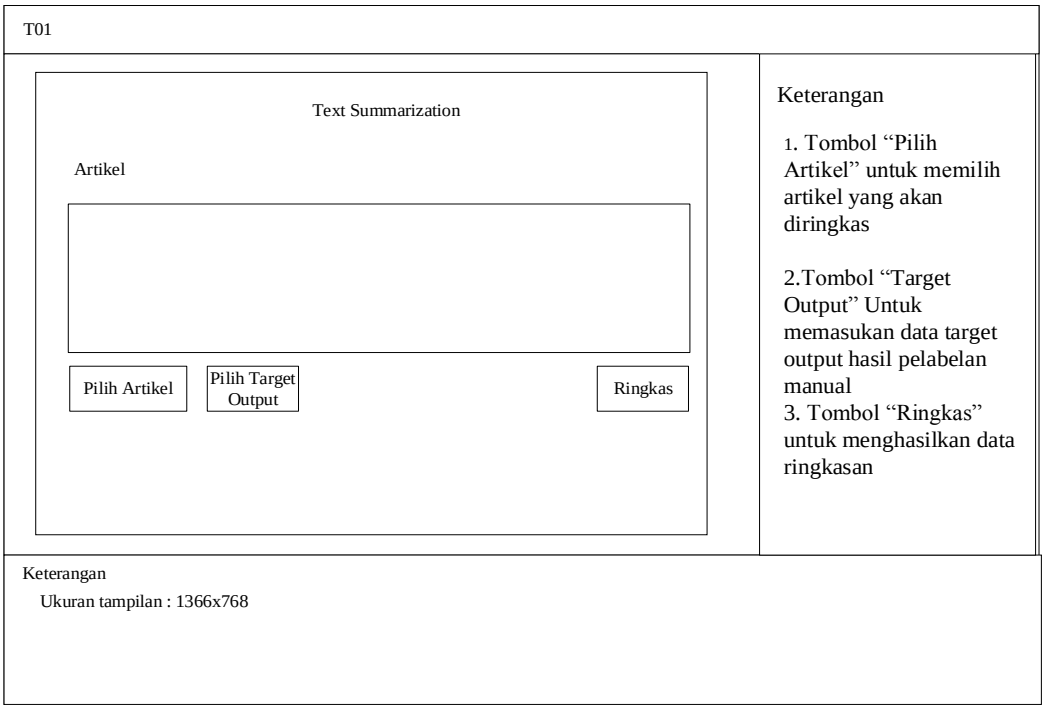
Gambar 3. 26 Perancangan Struktur Menu

3.6 Perancangan Antarmuka

Perancangan antarmuka menjelaskan tentang rencana antarmuka yang digunakan pada sistem peringkasan teks otomatis ini. Adapun rancangan antarmuka-nya adalah sebagai berikut.

3.6.1 Perancangan Antarmuka Halaman Utama

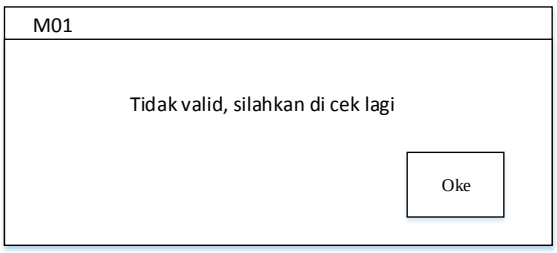
Perancangan antar muka halaman utama adalah halaman utama pada perancangan peringkasan teks otomatis. Adapun perancangan antarmuka-nya pada gambar 3.27 sebagai berikut.



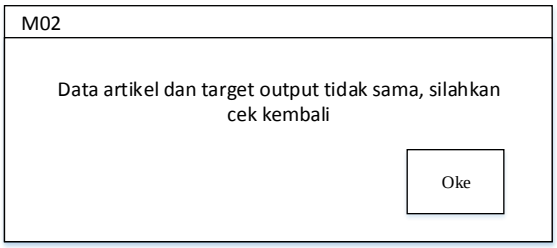
Gambar 3.27 Perancangan Antarmuka Halaman Utama

3.6.2 Perancangan Antarmuka Pesan Kesalahan

Perancangan antar muka pesan kesalahan merupakan pesan untuk input berita ataupun target *output* kepada pengguna jika belum memilih artikel berita atau target *output* dan artikel berita atau target *output* tidak sama pada sistem peringkasan teks otomatis. Adapun perancangan antarmuka-nya pada gambar 3.28 dan 3.29 sebagai berikut.



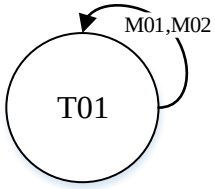
Gambar 3.28 Perancangan Antarmuka Pesan Kesalahan Data Kosong



Gambar 3.29 Perancangan Antarmuka Pesan Kesalahan Data Tidak Sama

3.7 Perancangan Jaringan Semantik

Jaringan semantik menggambarkan hubungan antara menu atau antarmuka dengan antarmuka lainnya. Adapun perancangan jaringan semantik untuk sistem yang dibangun dapat dilihat pada gambar 3.30 sebagai berikut.



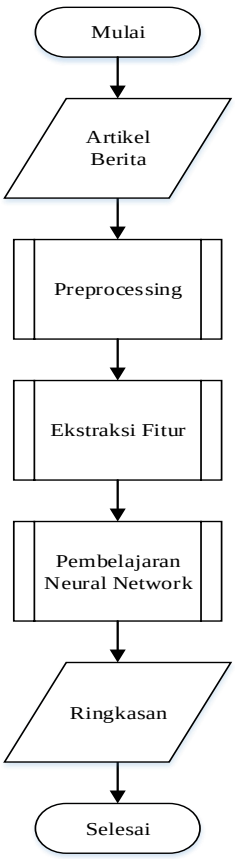
Gambar 3.30 Perancangan Jaringan Semantik

3.8 Perancangan Prosedural

Perancangan prosedural dalam penelitian ini diantaranya sebagai berikut.

3.8.1 Perancangan Algoritma Utama

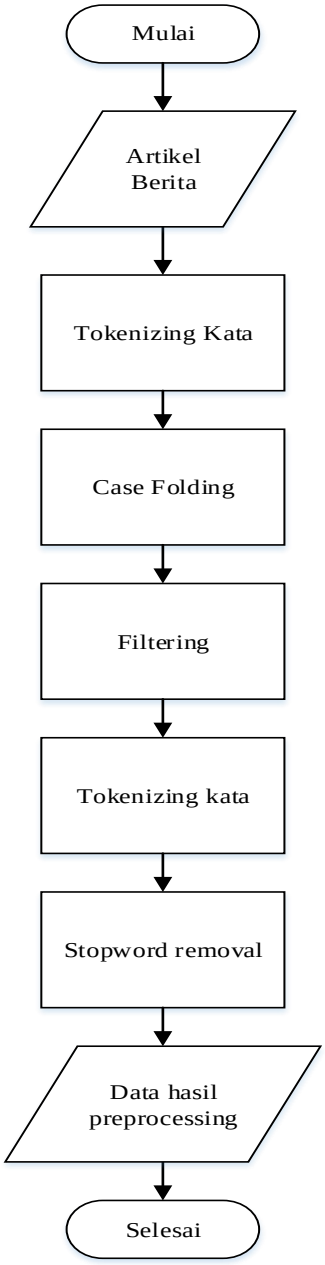
Perancangan algoritma utama dalam langkah atau proses peringkasan teks otomatis dengan metode *neural network* pada artikel Bahasa Indonesia ini dapat dilihat pada gambar 3.31 sebagai berikut



Gambar 3.31 *Flowchart* Algoritma Utama

3.8.2 Perancangan Algoritma *Preprocessing*

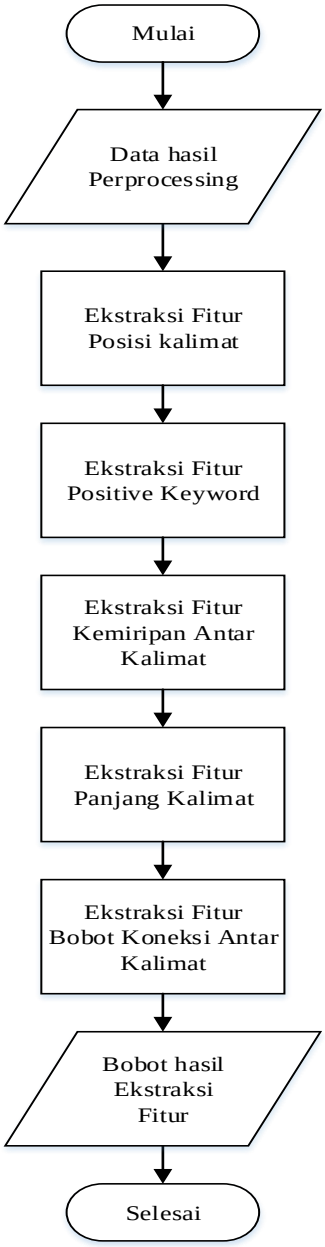
Perancangan *preprocessing* merupakan tahapan pengelolaan kata pada kalimat yang akan digunakan pada proses selanjutnya ekstraksi fitur. *Preprocessing* meliputi *tokenizing* kalimat, *case folding*, *filtering*, *tokenizing* kata dan *stopword removal*. Berikut ini merupakan perancangan *preprocessing* dapat dilihat pada gambar 3.32.



Gambar 3.32 *Flowchart* Algoritma *Preprocessing*

3.8.3 Perancangan Algoritma Ekstraksi Fitur

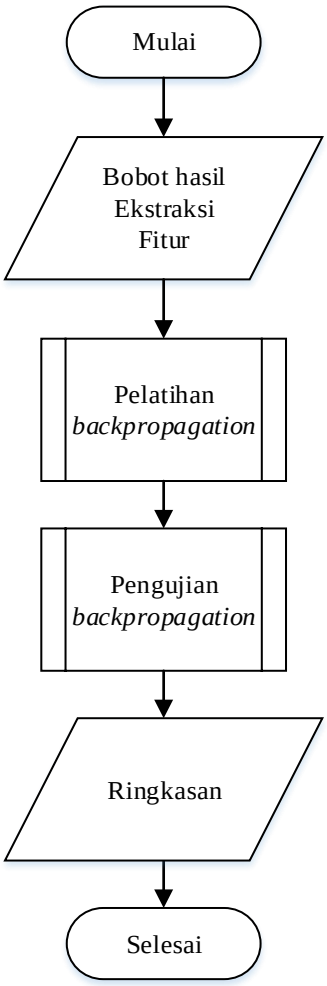
Ekstraksi fitur merupakan proses pengelolaan kalimat untuk menghasilkan nilai-nilai bobot. Nilai bobot hasil ekstraksi fitur merupakan masukan yang digunakan pada metode pembelajaran *neural network*. Fitur ekstraksi yang digunakan adalah posisi kalimat, *positive keyword*, kemiripan antar kalimat, panjang kalimat dan bobot koneksi antar kalimat. Perancangan ekstraksi fitur dapat dilihat pada gambar berikut 3.33 sebagai berikut.



Gambar 3.33 *Flowchart* Algoritma Ekstraksi Fitur

3.8.4 Perancangan Algoritma Pembelajaran *Neural Network*

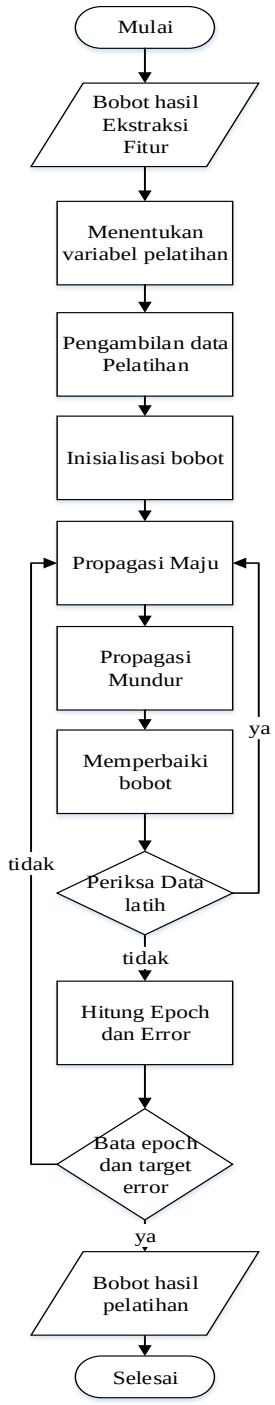
Pembelajaran *neural network* merupakan tahapan pengelolaan data hasil bobot ekstraksi fitur yang sebelumnya telah dilakukan. *Neural network* yang digunakan adalah *backpropagation*. Proses pembelajaran terdiri dari dua yaitu pelatihan *backpropagation* dan pengujian *backpropagation*. Perancangan pembelajaran *neural network* dapat dilihat pada gambar berikut 3.34 sebagai berikut



Gambar 3.34 *Flowchart* Algoritma *Case Folding*

3.8.5 Perancangan Algoritma Pelatihan *Backpropagation*

Pelatihan *backpropagation* merupakan proses pertama dalam pembelajaran *neural network*. Data input masukan dari proses ini adalah data hasil bobot ekstraksi fitur. Proses pelatihan *backpropagation* mengelola data yang telah dimasukan berdasarkan kalimat yang dimana bobot ekstraksi fitur yang terdiri dari lima ekstraksi dijadikan input masukan pada input layer, proses dilanjutkan dengan pemeriksaan data latih dan proses akan berhenti jika nilai *epoch* atau target *error* telah terpenuhi. Perancangan pelatihan *backpropagation* dapat dilihat pada gambar berikut 3.35 sebagai berikut

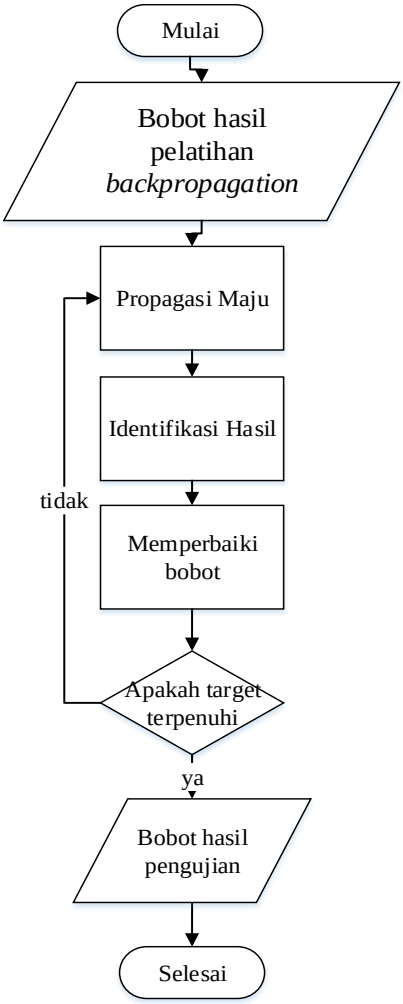


Gambar 3. 35 *Flowchart* Algoritma Pelatihan *Backpropagation*

3.8.6 Perancangan Algoritma Pengujian *Backpropagation*

Pengujian *backpropagation* merupakan proses pengecekan dengan melalui tahapan propagasi maju yang dimana proses ini akan mengecek bobot yang telah dihasilkan dari proses pelatihan *backpropagation* apakah bobot yang dihasilkan

sesuai dengan target yang ditentukan. Perancangan pengujian *backpropagation* dapat dilihat pada gambar berikut 3.36 sebagai berikut



Gambar 3.36 *Flowchart* Algoritma Pengujian *Backpropagation*