

BAB IV

ANALISA

4.1. Analisa Data

Pada penelitian ini data yang digunakan bersumber dari postingan pelanggan pada akun Facebook Group iRaise Helpdesk, akun Facebook Group tersebut berfungsi sebagai layanan pelanggan (*customer service*) terhadap sistem iRaise pada Universitas Islam Negeri Sultan Syarif Kasim Riau (UIN Suska Riau). Yang saat ini telah beranggotakan ± 11.605 anggota.

Berikut merupakan tampilan halaman pada akun Facebook Group iRaise Helpdesk UIN Suska Riau seperti Gambar 4.1 dibawah ini:



Gambar 4. 1 Halaman Facebook Group iRaise Helpdesk

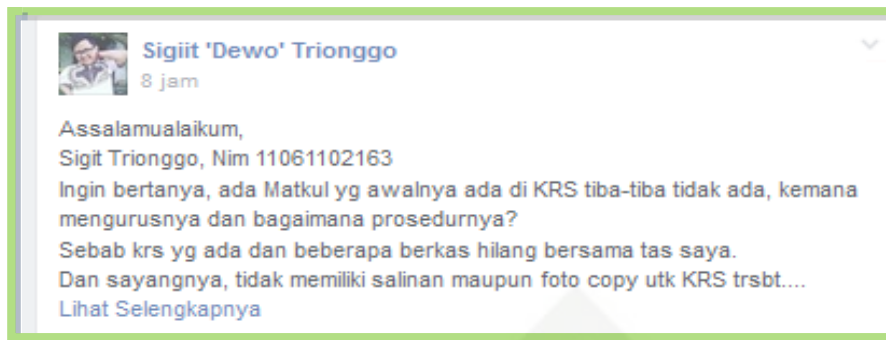
Pada gambar bertabel diatas menunjukan bahwa pada gambar petunjuk 1 merupakan postingan keluhan pelanggan terhadap sistem iRaise, dan pada gambar petunjuk 2 merupakan sebuah komentar pelanggan terhadap postingan. Sehingga pada penelitian ini data yang digunakan ialah data yang hanya berdasarkan postingan pelanggan mengenai keluhan terhadap sistem iRaise seperti gambar petunjuk 1. Dengan hasil Gambar 4.2 dibawah ini:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4. 2 Keluhan Sistem iRaise

4.1.1. Pengambilan Data

Data yang diambil dari akun Facebook Group iRaise Helpdesk adalah data berdasarkan postingan yang hanya mengandung kalimat keluhan mengenai permasalahan sistem iRaise yang di ambil secara manual, kemudian di simpan dalam bentuk *microsoft excel*. Seluruh data yang terkumpul berjumlah 1660 data postingan. Kemudian data tersebut dijadikan tabel baru yaitu tabel keluhan sistem iRaise dengan menambahkan kolom no, kolom keluhan, dan kategori secara manual. Pada kolom no akan diisi nomor urut keluhan sistem iRaise, kolom keluhan berisi kalimat keluhan sistem iRaise, dan pada kolom kategori diisikan label (kelas) dari keluhan sistem iRaise yang juga ditambahkan secara manual. Berikut Tabel 4.1 yang menjelaskan bagian-bagian keluhan sistem iRaise:

Tabel 4. 1 Deskripsi Tabel Keluhan Sistem iRaise

Nama kolom	Deskripsi
No	Nomor urut
Keluhan	Isi dari keluhan sistem iRaise
Kategori	Label dari keluhan (Login, KRS, Nilai, dan Personal)

4.1.2. Pelabelan Data

Berdasarkan pada landasan teori tentang keluhan maka pada penelitian ini pelabelan terbagi dalam empat kelas kategori yaitu kelas login, krs, nilai dan kelas personal. Pelabelan pada penelitian ini dilakukan secara manual, di kelaskan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

berdasarkan ciri-ciri kategorinya. Ciri-ciri yang termasuk dalam kelas kategori adalah sebagai berikut:

Tabel 4. 2 Ciri-ciri Keluhan Pada Kelas Kategori

Kelas Kategori	Ciri-ciri Keluhan
Login	<ul style="list-style-type: none"> - Kesalahan password atau kesalahan id - Kesalahan lupa password - Kesalahan email atau lupa password email - Pembayaran spp → ex : ketika login muncul anda belum melakukan pembayaran spp.
Krs	<ul style="list-style-type: none"> - Kesalahan terhadap Kartu Rencana Studi (KRS) - Kesalahan terhadap sks - Kesalahan ambil matakuliah atau hapus matakuliah - Tidak dapat mengambil matakuliah (kelas penuh) - Jadwal bentrok atau jadwal pada krs berubah - Perubahan dosen pembimbing akademik (PA) - Kesalahan persetujuan krs → PA
Nilai	<ul style="list-style-type: none"> - Kesalahan nilai - Kesalahan pada kartu hasil studi (khs) - Kesalahan terhadap transkrip nilai
Personal	<ul style="list-style-type: none"> - Audit kinerja dosen - Edit profil - Cuti kuliah

Jumlah data yang telah dikumpulkan diberi label berdasarkan kelasnya berjumlah 1660 data postingan, yang terdiri dari 770 kelas login, 260 kelas nilai, 370 kelas krs, dan 260 kelas personal. Kemudian dataset tersebut dilakukan pemotongan menjadi *balanced* dataset berdasarkan jumlah data minoritas dengan hasil data berjumlah 1040 data yang berisi 260 data pada masing-masing kelas keluhan, yaitu 260 pada kelas login, 260 kelas krs, 260 nilai dan 260 kelas personal. Proses penerapan *balanced* dataset dilakukan untuk mencegah proses penurunan tingkat akurasi ketika di terapkan pada dataset dengan *imbalance*, *ratio* yang tinggi seringkali memprediksi semua *instance* sebagai kelas yang menjadi mayoritas dalam pelatihan, dan data dari kelas minoritas yang mungkin dianggap *noise* sehingga diabaikan oleh teknik klasifikasi. Jumlah dataset tersebut di implementasikan dalam grafik berikut ini:

Hak Cipta Dilindungi Undang-Undang

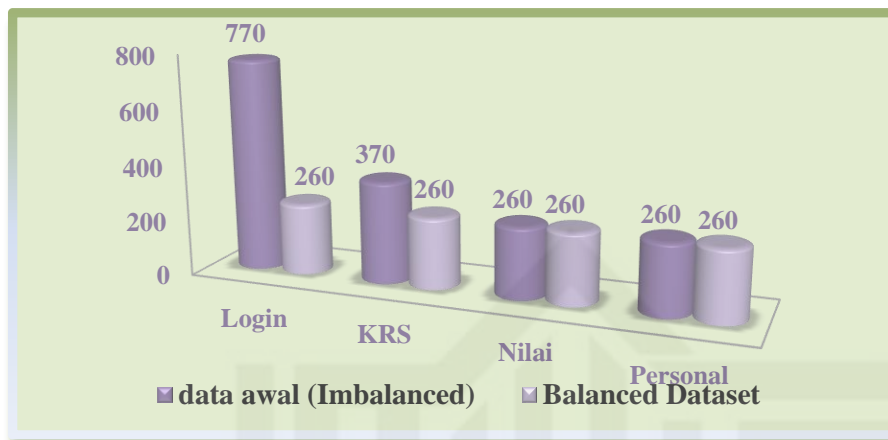
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Data Keluhan Sistem iRaise Berlabel



Gambar 4. 3 Grafik Data keluhan Sistem iRaise

Berikut hasil data keluhan sistem iRaise yang telah diberi label dapat dilihat pada Tabel 4.3 bawah ini:

Tabel 4. 3 Hasil Label Keluhan Sistem iRaise

No	Keluhan	Kategori
1	Nama:Riswan Hasibuan, NIM: 11521105581. E-mail: Riswan.hasber05@gmail.com. kenapa iRaise saya tidak bisa dibuka pak/bu? Selalu saja muncul tulisan error. mohon bantuannya.	Login
2	Assalamualaikum,Mohon bantuan nya pak/ibu, pas saya mengisi krs untuk smster 4 kok mata kuliah nya gx ada timbul satu pun Mohon bantuan nya pak/ibu Wassalamualaikum	KRS
3	nama: aila risda wati, fak: fekonsos, jurusan: akuntansi, saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya. Mohon bantuannya.	Personal
4	Mw tanya pak/buk nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol. mohon solusinya pak/buk. Karena saya takut nilai saya hilang	Nilai
5	permasalahan krs saya tidak bisa diisi, pada saat isi, selalu di bilang melebihi kuota. mohon sekali bantuan dari bapak/ibu.	KRS
6	saya tidak pernah mengajukan cuti tetapi di iraise saya mengajukan cuti. Jadi saya mohon solusi nya bapak/ibuk	Personal

...		
1040	Assalamu'alaikum admin. saya mau bertanya, knapa sewaktu saya login ke iraise, ada pemberitahuan pengajuan cuti anda sedang di proses, sedangkan saya tidak pernah mengajukan cuti. mohon bantuannya min	Personal

4.1.3. Validasi Pelabelan Data

Pada penelitian ini dilakukan proses validasi terhadap hasil pelabelan data secara manual, proses validasi dilakukan dengan melakukan uji perbandingan antara label data manual dengan label data hasil kuesioner. Hasil pelabelan data kuesioner dilakukan dengan beberapa *sample* data keluhan terhadap sistem iRaise, data kuesioner diberikan 25 *sample* data keluhan yang di tujukan kepada 8 responden yaitu admin iRaise Helpdesk UIN Suska Riau.

Berikut merupakan contoh *sample* data kuesioner beserta jawaban responden terhadap keluhan sistem iRaise dapat dilihat pada Tabel 4.4 dibawah ini:

Tabel 4. 4 Sample Data Kuesioner

No	Keluhan	Jawaban Responden
1	kenapa iRaise saya tidak bisa dibuka pak/bu? Selalu saja muncul tulisan error. mohon bantuannya.	8 Responden Login
2	Assalamualaikum,Mohon bantuan nya pak/ibu, pas saya mengisi krs untuk smster 4 kok mata kuliah nya gx ada timbul satu pun Mohon bantuan nya pak/ibu Wassalamualaikum	8 Responden KRS
3	saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya. Mohon bantuannya.	8 Responden Personal
4	Mw tanya pak/buk nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol. mohon solusinya pak/buk. Karena saya takut nilai saya hilang	8 Responden Nilai
5	permasalahan krs saya tidak bisa diisi, pada saat isi, selalu di bilang melebihi	8 Responden

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	kuota. mohon sekali bantuan dari bapak/ibu.	KRS
13	Saya gak bisa liat nilai di iraise, Setiap kali masuk pasti isi audit dosen terus, setiap dah di simpan katanya terimakasih, ikuti kousener swlanjutnya, setiap saya ikuti kyak gt trus gk berhnti” ngulang dr awal terus, bantu jawab kak	7 Responden Personal dan 1 Nilai.
...		
24	Pak/ibu, knapa saya tidak bisa mengisi krs. Setiap kali saya isi krs, yg muncul selalu profil.	6 Responden Personal, 1 KRS, dan 1 Nilai.
25	Saya mau tanya, tadi saya isi krs, tapi saya salah mengambil 1 mata kuliah, padahal sebelumnya sudah saya cek betul, jadi apakah bisa dirubah lagi? Terimakasih sebelumnya	8 Responden KRS

Berikut merupakan perbandingan hasil dari pelabelan secara manual dengan pelabelan berdasarkan responden admin iRaise Helpdesk UIN Suska Riau.

Tabel 4. 5 Perbandingan Hasil Label

Keluhan	Label Manual	Label Responden
Keluhan 1	Login	Login
Keluhan 2	KRS	KRS
Keluhan 3	Personal	Personal
Keluhan 4	Nilai	Nilai
Keluhan 5	KRS	KRS
Keluhan 13	Personal	Personal
:	:	:
Keluhan 24	Personal	Personal
Keluhan 25	KRS	KRS

Berdasarkan Tabel 4.4 dan Tabel 4.5 dapat dilihat perbandingan hasil dari pelabelan data keluhan secara manual dengan pelabelan responden admin iRaise Helpdesk UIN Suska Riau, kedua nya memiliki jawaban yang sama dan cukup signifikan sebesar 98% dan signifikansi sebesar 2%.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

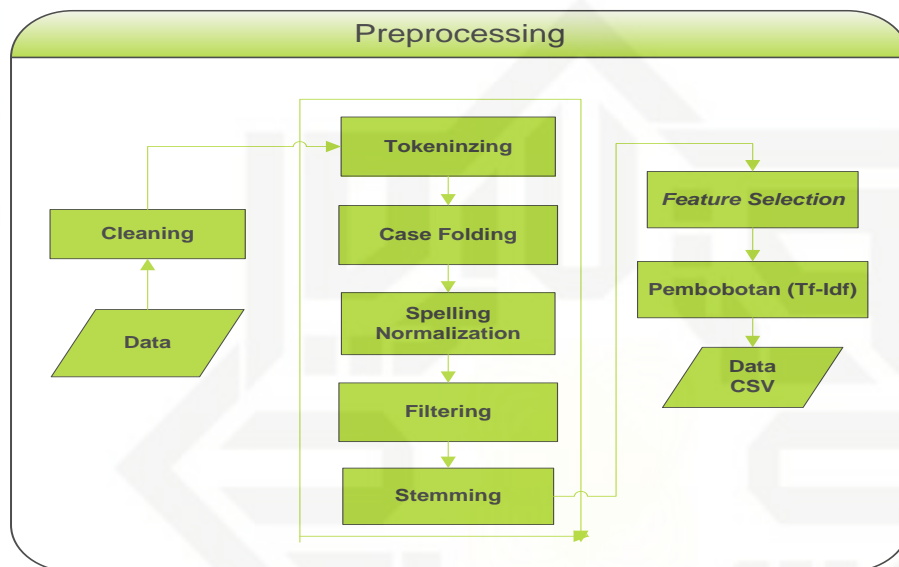
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.2. Text Preprocessing

Text Preprocessing merupakan langkah penting dalam melakukan analisa klasifikasi keluhan yang bertujuan untuk membersihkan data dari unsur-unsur yang ada yang tidak dibutuhkan. Gambar 4.4 berikut ini merupakan langkah secara umum *preprocessing text*:



Gambar 4. 4 Tahapan *Text Preprocessing*

Berikut ini penjelasan Gambar 4.4 yang merupakan tahapan *text preprocessing* sebagai berikut:

1. Cleaning

Adapun kata atau karakter yang dihilangkan adalah identitas pengguna seperti: alamat url (*email.com*), *username* (nama pengguna sistem iRaise), *nim* (nomor induk mahasiswa), *fakultas* (nama fakultas pengguna), *jurusan* (jurusan pengguna). Proses *cleaning* dengan menghilangkan identitas pengguna dilakukan untuk membersihkan data dari kata yang tidak penting dan tidak mempengaruhi proses klasifikasi. Proses *cleaning* dilakukan sekaligus pada saat pengumpulan data. Hasil *cleaning* dari data keluhan sistem iRaise dapat dilihat pada Tabel 4.6 adalah sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 4. 6 Hasil *Cleaning*

No	Keluhan	Kategori
1	kenapa iRaise saya tidak bisa dibuka pak/bu? Selalu saja muncul tulisan error. mohon bantuannya.	Login
2	Assalamualaikum, Mohon bantuan nya pak/ibu,,pas saya mengisi krs untuk smster 4 kok mata kuliah nya gx ada timbul satu pun Mohon bantuan nya pak/ibu Wassalamualaikum	KRS
3	saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya. Mohon bantuannya.	Personal
4	Mw tanya pak/buk nilai saya sudah ada beberapa yg keluar tpi bobot nya masih nol. mohon solusinya pak/buk. Karena saya takut nilai saya hilang	Nilai
5	permasalahan krs saya tidak bisa diisi, pada saat isi, selalu di bilang melebihi kuota. mohon sekali bantuan dari bapak/ibu.	KRS
6	saya tidak pernah mengajukan cuti tetapi di iraise saya mengajukan cuti. Jadi saya mohon solusi nya bapak/ibuk	Personal

2. *Tokenizing*

Tokenizing yaitu proses melakukan pemecahan terhadap kalimat data keluhan menjadi sebuah kata tunggal. Pada saat bersamaan, tokenisasi juga membuang beberapa angka, karakter tertentu yang dianggap sebagai tanda baca. Berikut merupakan proses *tokenizing* untuk contoh keluhan sistem iRaise dapat di lihat pada Tabel 4.7 berikut ini:

Tabel 4. 7 Hasil *Tokenizing*

Transform case	Tokenizing	Transform case	Tokenizing
Data 1 : kenapa iRaise saya tidak bisa dibuka	kenapa	Data 2 : Assalamualaik um, Mohon bantuan nya	Assalamualaikum
	iRaise		Mohon
	saya		bantuan
	tidak		nya

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

pak/bu?	bisa	pak/ibu,,pas	pak
Selalu saja	dibuka	saya mengisi	ibu
muncul	pak	krs untuk	pas
tulisan error.	bu	smster 4 kok	saya
mohon	Selalu	mata kuliah	mengisi
bantuannya.	saja	nya gx ada	krs
	muncul	timbul satu	untuk
	tulisan	pun Mohon	smster
	error	bantuan nya	kok
	mohon	pak/ibu	matakuliah
	bantuannya	Wassalamualai	nya
		kum	gx
	saya		ada
	setiap		timbul
	ngisi		satu
	audit		pun
	selalu		Mohon
	kembali		bantuan
	ke		nya
	audit		pak
	yang		ibu
	sudah		Wassalamualaiku
	di		m
Data 3: saya setiap ngisi audit, selalu kembali ke audit yang sudah di isi, tidak ke audit selanjutnya. Mohon bantuannya.	isi	Data 4 :	Mw
	tidak	Mw tanya	tanya
	ke	pak/buk nilai	pak
	audit	saya sudah ada	buk
	selanjutnya	beberapa yg	nilai
	Mohon	keluar tpi	saya

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	bantuannya	bobot nya	sudah
		masih nol.	ada
	permasalahan	mohon	beberapa
Data 5 :	krs	solusinya	yg
permasalahan	saya	pak/buk.	keluar
krs saya tidak	tidak	Karena saya	tpi
bisa diisi,	bisa	takut nilai saya	bobot
pada saat isi,	diisi	hilang	nya
selalu di	pada		masih
bilang	saat		nol
melebihi	isi		mohon
kuota. mohon	selalu		solusinya
sekali	di		pak
bantuan dari	bilang		buk
bapak/ibu.	melebihi		Karena
	kuota		saya
	mohon		takut
	sekali		nilai
	bantuan		saya
	dari		hilang
	bapak		
	Ibu		
Data 6:	saya	Di	saya
saya tidak	tidak	Iraise	mohon
pernah	pernah	Saya	solusi
mengajukan	mengajukan	Mengajukan	nya
cuti tetapi di	cuti	Cuti	bapak
iraise saya	tetapi	Jadi	ibuk
mengajukan			
cuti. Jadi			
saya mohon			
solusi nya			

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. Case folding

Case folding merupakan proses penyeragaman bentuk huruf pada dokumen dengan mengubah huruf besar menjadi huruf kecil atau sebaliknya.

Tabel 4.8 berikut adalah hasil keluhan yang telah dilakukan proses *case folding*.

Tabel 4.8 Hasil Case folding

Keluhan 1	Keluhan 2	Keluhan 3
kenapa	assalamualaikum	saya
iraise	mohon	setiap
saya	bantuan	mengisi
tidak	nya	audit
bisa	bapak	selalu
dibuka	ibu	kembali
pak	pas	ke
buk	saya	audit
selalu	mengisi	yang
saja	krs	sudah
muncul	untuk	di
tulisan	smester	isi
error	kok	tidak
mohon	matakuliah	ke
bantuannya	nya	audit
Keluhan 4	gk	selanjutnya
mw	ada	mohon
tanya	timbul	bantuannya
pak	satupun	Keluhan 6
buk	pun	saya
nilai	mohon	tidak
saya	bantuan	pernah
sudah	nya	mengajukan
ada	bapak	cuti

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

beberapa	ibu	tetapi
yg	wassalamualaikum	di
keluar	Keluhan 5	iraise
tpi	permasalahan	saya
bobot	krs	mengajukan
nya	saya	cuti
masih	tidak	jadi
nol	bisa	saya
mohon	diisi	mohon
solusinya	pada	solusi
pak	saat	nya
buk	isi	bapak
karena	selalu	ibuk
saya	di	
takut	bilang	
nilai	melebihi	
saya	kuota	
hilang	mohon	
	sekali	
	bantuan	

4. Spelling Normalization

Merupakan proses perbaikan kata yang tidak sesuai dengan penulisan kata yang sebenarnya misalnya “gk” diubah menjadi “tidak”. Dapat dilihat pada Tabel 4.9 berikut:

Tabel 4. 9 Hasil Spelling Normalization

Keluhan 1	Keluhan 2	Keluhan 3
kenapa	assalamualaikum	saya
iraise	mohon	setiap
saya	bantuan	mengisi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tidak	nya	audit
bisa	bapak	selalu
dibuka	ibu	kembali
bapak:pak	ketika:pas	ke
ibu:buk	saya	audit
selalu	mengisi	yang
saja	krs	sudah
timbul:muncul	untuk	di
tulisan	semester:smester	isi
rusak:error	kok	tidak
mohon	matakuliah	ke
bantuannya	nya	audit
Keluhan 4	tidak:gk	selanjutnya
ingin:mw	ada	mohon
tanya	timbul	bantuannya
bapak:pak	satupun	Keluhan 6
ibu:buk	pun	saya
nilai	mohon	tidak
saya	bantuan	pernah
sudah	nya	mengajukan
ada	bapak	cuti
beberapa	ibu	tetapi
yang:yg	wassalamualaikum	di
keluar	Keluhan 5	iraise
tapi:tpi	permasalahan	saya
bobot	krs	mengajukan
nya	saya	cuti
masih	tidak	jadi
nol	bisa	saya
mohon	diisi	mohon

solusinya	pada	solusi
bapak:pak	saat	nya
ibu:buk	isi	bapak
karena	selalu	ibu:ibuk
saya	di	
takut	bilang	
nilai	melebihi	
saya	kuota	
hilang	mohon	
	sekali	
	bantuan	

5. Filtering

Filtering adalah tahap mengambil kata-kata penting dari hasil token. Pada tahap ini kata-kata yang bersifat *stopword* akan dibuang dikarenakan *stopword* dianggap tidak berguna, dimana kata-kata yang tidak digunakan akan dihilangkan dari data, menggunakan algoritma *stop-list* (membuang kata-kata kurang penting) atau *word-list* (menyimpan kata penting). Kata-kata yang tidak penting dikumpulkan dan dimasukkan kedalam kamus *stopword*. Hasil *filtering* dapat dilihat pada Tabel 4.10 dan contoh kamus *stopword* dapat dilihat pada Tabel 4.11 berikut:

Tabel 4. 10 Hasil Filtering

Filtering			
iraise	matakuliah	keluar	mengajukan
tidak	satupun	tetapi	cuti
bisa	setiap	bobot	diiraise
dibuka	mengisi	masih	timbul
selalu	audit	nol	login
timbul	selalu	takut	salah
tulisan	kembali	hilang	ipk
rusak	sudah	permasalahan	khs
mengisi	isi	diisi	terakhir
krs	selanjutnya	melebihi	
semester	nilai	kuota	

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

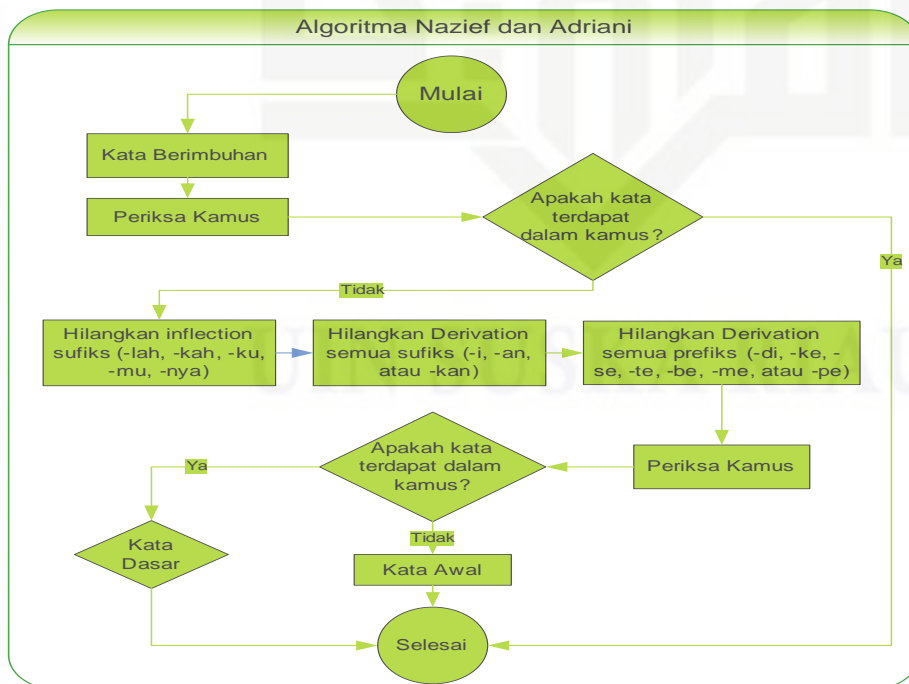
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 4. 11 Contoh Kamus *Stopwords*

Stopwords		
kenapa	assalamualaikum	dari
mohon	ke	untuk
saya	yang	saat
nya	di	karena
bapak	masih	sekali
ibu	ada	saja
ketika	wassalamualaikum	ingin
jadi	kok	tanya

6. *Stemming*

Pada tahap ini kata berimbuhan akan diubah dalam bentuk kata dasar dengan menggunakan algoritma Nazief dan Adriani berdasarkan Kamus Besar Bahasa Indonesia (KBBI). Proses *stemming* dengan algoritma Nazief dan Adriani dapat dilihat pada Gambar 4.5 *flowchart* berikut:



Gambar 4. 5Flowchat Algoritma Nazief dan Adriani

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Berikut contoh hasil *stemming* dapat di lihat pada Tabel 4.12 berikut:

Tabel 4. 12 Proses *Stemming*

<i>Stemming</i>		
buka:dibuka	lanjut:selanjutnya	isi:diisi
lalu:selalu	luar:keluar	lebih:melebihi
tulis:tulisan	tapi:tetapi	aju:mengajukan
isi:mengisi	masalah:permasalahan	iraise:diiraise
semester:smester	hilang:hilangkan	lanjut:lanjutkan

Tabel 4. 13 Daftar Kata Baku

Kata Baku				
iraise	tidak	bisa	buka	lalu
isi	krs	semester	matakuliah	timbul
tiap	audit	kembali	udah	lanjut
bobot	tulis	rusak	nol	cuti
nilai	luar	tapi	masalah	takut

7. Ekstrasi *Feature*

Pada penelitian ini pemilihan *feature selection* dilakukan dengan 3 cara ekstraksi *feature* yaitu pemilihan *feature* ekstraksi *threshold*, *ranking selection threshold up*. *Ranking selection threshold down*. pemilihan *feature* ekstraksi *threshold* adalah pengamatan terhadap distribusi frekuensi kemunculan kata dan jumlah *feature*, nilai *threshold* terbaik adalah titik dimana frekuensi kemunculan kata dan jumlah *feature* mulai konstan. Pemilihan *feature ranking selection* adalah pemberian *ranking* pada setiap *feature* yang ada dan mengesampingkan *feature* yang tidak memenuhi *standart* tertentu, *ranking selection* menentukan tingkat *ranking* secara independen antara satu *feature* dengan *feature* yang lainnya. *Feature ranking selection threshold up* merupakan *feature* yang mempunyai *ranking* tertinggi akan digunakan dan yang terendah akan dikesampingkan. *Feature ranking selection threshold down* merupakan *feature* yang mempunyai *ranking* terendah akan digunakan dan yang tertinggi akan dikesampingkan. Proses pemilihan *feature* ini yang nantinya akan di gunakan pda saat proses pengujian untuk melihat tingkat keakuratan klasifikasi pada dataset.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pemilihan *feature* ini dilakukan dengan cara merangking nilai kemunculan kata. Berikut Tabel 4.14 Ekstraksi *Feature*.

Tabel 4. 14 Ekstraksi *Feature*

Frekuensi Kemunculan <i>Feature</i>	Jumlah <i>Feature</i>
1	590
2	399
3	315
4	274
5	243
6	215
7	200
8	186
9	171
10	157
11	152
12	149
13	142
14	139
15	135
16	131
17	125
18	121
19	120
20	116
21	110
22	108
23	105
24	100
≥ 25	97

Berdasarkan tabel dari pengamatan terhadap nilai *threshold*, jumlah *feature* yang dihasilkan mulai konstan pada saat nilai *threshold* berada pada kisaran 13-24.

8. Pembobotan

Data yang sudah melalui proses *text preprocessing* kemudian akan diberi bobot dengan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Pembobotan itu sendiri merupakan proses merubah kata menjadi bentuk *vektor*, dimana tiap kata dihitung sebagai satu fitur. Dengan melakukan beberapa proses yaitu menghitung *Term Frequency* (TF), *Document Frequency* (DF), *Inverse Document Frequency* (IDF) dan mengalikan TF dengan IDF sebagai bentuk bobot dari kata suatu *keyword*.

Untuk mencari nilai IDF dapat menggunakan rumus 2.2 berikut adalah contoh perhitungan IDF dengan menggunakan kata “iraise”.

$$IDF(iraise) = \log \frac{6}{2} = 0.47712$$

Berikut adalah contoh perhitungan bobot *Term Frequency-Inverse Document Frequency* (TF-IDF) menggunakan rumus 2.1

$$w(iraise) = 1 \cdot 0.47712 = 0,477$$

Pada Tabel 4.15 berikut merupakan contoh hasil transformasi berdasarkan contoh keluhan diatas.

Tabel 4. 15 Penghitungan TF dan IDF

Kata	TF						DF	IDF
	data 1	data 2	data 3	data 4	data 5	data 6		=log (n/df) n=6
iraise	1	0	0	0	0	1	2	0,47712
tidak	1	1	1	0	1	1	5	0.07918
bisa	1	0	0	0	1	0	2	0,47712
buka	1	0	0	0	0	0	1	0.77815
isi	0	1	2	0	2	0	3	0.30102
krs	0	1	0	0	1	0	2	0,47712
semester	0	1	0	0	0	0	1	0.77815
matakuliah	0	1	0	0	0	0	1	0.77815
timbul	1	1	0	0	0	0	2	0,47712
tiap	0	0	1	0	0	0	1	0.77815
audit	0	0	3	0	0	0	1	0.77815
kembali	0	0	1	0	0	0	1	0.77815
sudah	0	0	1	1	0	0	2	0,47712

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

lanjut	0	0	1	0	0	0	1	0.77815
hilang	0	0	0	1	0	0	1	0.77815
masalah	0	0	0	0	1	0	1	0.77815
lebih	0	0	0	0	1	0	1	0.77815
kuota	0	0	0	0	1	0	1	0.77815
tulis	1	0	0	0	0	0	1	0.77815
rusak	1	0	0	0	0	0	1	0.77815
cuti	0	0	0	0	0	2	1	0.77815
nilai	0	0	0	2	0	0	1	0.77815
luar	0	0	0	1	0	0	1	0.77815
tapi	0	0	0	0	0	1	1	0.77815
nol	0	0	0	1	0	0	1	0.77815
takut	0	0	0	1	0	0	1	0.77815
bobot	0	0	0	1	0	0	1	0.77815

Tabel 4. 16 Penghitungan TF-IDF

Kata	TF-IDF=(TFxIDF)					
	data 1	data 2	data 3	data 4	data 5	data 6
Iraise	0,477	0	0	0	0	0,477
Tidak	0.079	0.079	0.079	0	0.079	0.079
Bisa	0,477	0	0	0	0,477	0
Buka	0.778	0	0	0	0	0
Isi	0	0.301	0.602	0	0.602	0
Krs	0	0,477	0	0	0,477	0
semester	0	0.778	0	0	0	0
matakuliah	0	0.778	0	0	0	0
Timbul	0,477	0,477	0	0	0	0
Tiap	0	0	0.778	0	0	0
Audit	0	0	2.334	0	0	0
kembali	0	0	0.778	0	0	0
Sudah	0	0	0,477	0,477	0	0
Lanjut	0	0	0.778	0	0	0
Hilang	0	0	0	0.778	0	0
masalah	0	0	0	0	0.778	0
Lebih	0	0	0	0	0.778	0
Kuota	0	0	0	0	0.778	0
Tulis	0.778	0	0	0	0	0
rusak	0.778	0	0	0	0	0
Cuti	0	0	0	0	0	1.556
Nilai	0	0	0	1.556	0	0
Luar	0	0	0	0.778	0	0
Tapi	0	0	0	0	0	0.778
Nol	0	0	0	0.778	0	0
Takut	0	0	0	0.778	0	0
Bobot	0	0	0	0.778	0	0

4.3. Proses SVM

Pemodelan pada penelitian ini menggunakan metode *Support Vector Machine* (SVM) dalam melakukan klasifikasi data. Proses klasifikasi akan menggunakan *tools* RapidMiner pada implementasi nantinya. Di dalam proses

pengklasifikasian terdapat proses pembelajaran (*training*) dengan menggunakan data latih dan menghasilkan model yang akan dilakukan proses pengujian (*testing*) dengan beberapa data yang akan digunakan sebagai data uji.

Berikut tahapan pemodelan SVM pada proses klasifikasi :

4.3.1. *Training*

Pada tahap ini *Machine Learning* akan diberikan data *training*. Data *training* terdiri dari data keluhan sistem iRaise yang sudah diberikan target label yang kemudian akan berfungsi memberikan pemahaman untuk mesin sehingga dapat mengklasifikasikan keluhan yang bernilai login, krs, nilai dan personal. Data *training* akan diproses menggunakan *function* LibSVM yang dapat digunakan untuk klasifikasi dan formula yang akan digunakan *C-Support Vector Classifier* (C-SVM) dengan kernel *Radial Basis Function* (RBF). Pada penelitian ini data yang digunakan berbentuk teks proses klasifikasi tidak dapat di asumsikan secara *linear*, oleh karena itu akan digunakan kernel sebagai parameter yang dapat melakukan klasifikasi terhadap data yang bersifat *multiclass* dengan menggunakan kernel RBF dengan mencari dua parameter c dan γ .

Parameter C merupakan parameter pengatur *Tradeoff* dari kesalahan pengklasifikasian data. Semakin kecil nilai C fungsi klasifikator yang terbentuk akan bersifat *soft margin*. Sedangkan jika nilai C semakin besar fungsi klasifikator yang terbentuk akan bersifat *hard margin*. Sedangkan nilai γ akan mempresentasikan tingkat kepentingan sebagai data *training* terhadap data *training* lainnya. Nilai yang besar berarti sebagian data *training* akan semakin berpengaruh pada bagian lain dari data *training*.

a. *Testing*

Proses *testing* digunakan untuk proses pengujian dataset untuk menemukan klasifikasi berdasarkan data *sample* pengujian. Klasifikasi merupakan suatu pengukuran yang dilakukan terhadap data *training* yang merupakan data latih, pada penelitian ini proses klasifikasi pada pembelajaran SVM diselesaikan dengan menggunakan *kernel trick* untuk data *non-linier*. Dengan menggunakan kernel *Gaussian RBF*.

Berikut contoh hitungan manual untuk mendapatkan hasil kernel sebagai berikut :

Tabel 4. 17 Data Keluhan Bernilai Vector

No	X ₁	X ₂	X ₃	X ₄	X ₅	...	X ₁₅	X ₁₆	Kelas
1	0.477	0	0	0	0		0	0	Login
2	0	0.301	0.477	0.778	0.778		0	0	KRS
3	0	0.602	0	0	0		0	0	Personal
4	0	0	0	0	0		0.778	0.778	Nilai
5	0	0.602	0.477	0	0		0	0	KRS
6	0.477	0	0	0	0		0	0	Personal

Agar data dapat diklasifikasi, maka data direpresentasikan sebagai *vektor* berupa pasangan data (x, γ) . Nilai x merupakan nilai *vektor* data keluhan. Nilai γ merupakan *class* prediksi diberi label untuk login diberi label 1, krs diberi label 2, nilai diberi label 3, dan *class* personal diberi label 4. Pasangan *vektor* dapat dilihat pada Tabel 4.18. berikut:

Tabel 4. 18 Pasangan Vector (x, γ)

No	X _{1,γ}	X _{2,γ}	X _{3,γ}	X _{4,γ}	X _{5,γ}	...	X _{15,γ}	X _{16,γ}
1	0.477,1	0	0	0	0	...	0	0
2	0	0.301,2	0.477,2	0.778,2	0.778,2	...	0	0
3	0	0	0	0	0	...	0.778,3	0.778,3
4	0	0.602,4	0	0	0	...	0	0
5	0	0.602,2	0.477,2	0	0	...	0	0
6	0.477,4	0	0	0	0	...	0	0

Berikut perhitungan *Kernel Gaussian RBF* :

$$\begin{aligned}
 X_1 - X_1' &= \begin{bmatrix} 0.477 & 0 & 0 & 0 & 0 & 0.477 & 0 & 0 & 0 & 0 & 0.778 & 0 & 0 & 0 & 0 \\ 0.477 & 0 & 0 & 0 & 0 & 0.477 & 0 & 0 & 0 & 0 & 0.778 & 0 & 0 & 0 & 0 \end{bmatrix} - \\
 &= [0] \\
 &= \exp\left(\frac{-||x-y||^2}{2\sigma^2}\right) = \frac{0}{2.2^2} = 0
 \end{aligned}$$

$$\begin{aligned}
 X_1 - X_2' &= \begin{bmatrix} 0.477 & 0 & 0 & 0 & 0 & 0.477 & 0 & 0 & 0 & 0 & 0.778 & 0 & 0 & 0 & 0 \\ 0.477 & 0.778 & 0.778 & 0.477 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} - \\
 &= [-0.301]
 \end{aligned}$$

$$= \exp\left(\frac{-||x-y||^2}{2\sigma^2}\right) = \frac{0.091}{2.2^2} = 0.0113$$

Selanjutnya dapat di lihat dalam Tabel 4.19 berikut :

Tabel 4. 19 Menghitung Kernel

x_i	$\exp\left(\frac{- x-y ^2}{2\sigma^2}\right)$
x_1-x_3'	0.238
x_1-x_4'	0.023
x_1-x_5'	0.002
x_1-x_6'	0.028
x_2-x_1'	0.0113
x_2-x_2'	0
x_2-x_3'	0.145
x_2-x_4'	0.002
x_2-x_5'	0.003
x_2-x_6'	0.075
:	
x_6-x_6'	0

Dengan melakukan perhitungan untuk seluruh data x_i maka didapat nilai matriks K sebagai berikut :

$$K = \begin{bmatrix} 0 & 0.011 & 0.238 & 0.022 & 0.002 & 0.028 \\ 0.0113 & 0 & 0.145 & 0.002 & 0.075 & 0.976 \\ 0.238 & 0.145 & 0 & 0.114 & 0.197 & 0.431 \\ 0.023 & 0.002 & 0.001 & 0 & 0.011 & 0.102 \\ 0.002 & 0.003 & 0.197 & 0.014 & 0 & 0.045 \\ 0.028 & 0.075 & 0.987 & 0.102 & 0.045 & 0 \end{bmatrix}$$

Matriksk sebagai pengganti $\text{dotproduct}x_i^T x_j$ pada persamaan $w = \sum_{i=1}^n \alpha_i y_i x_i$, pada persamaan *non linear separable data problem* dari formula pencarian *hyperplane* terbaik sulit dipecahkan maka digunakan bentuk dual yang merubah nilai w dalam bentuk α . Perubahan nilai w ke dalam bentuk α merupakan masalah *Quadratic Programming* (QP). Nilai α_i menunjukan *support vector* fungsi keputusan prediksi dari pelatihan. Sehingga pada penelitian ini *support vector* didapat dengan menggunakan *tools* RapidMiner.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.4. Analisa Evaluasi

Evaluasi model hasil proses SVM diukur dengan menghitung keakurasi data dengan proses pengujian terhadap data dengan menggunakan *cross-validation* dimana user harus menginput nilai *fold* yang akan digunakan untuk menemukan nilai akurasi terbaik. Nilai *default fold cross-validation* pada RapidMiner adalah 10, yang berarti bahwa data di bagi menjadi $K=10$ *fold*. Kemudian data tersebut akan dilakukan proses iterasi dimana satu data K akan dijadikan sebagai data uji dan selebihnya K akan dijadikan data latih. Proses ini akan dilakukan sampai seluruh data dapat menjadi data uji dan yang lain juga dapat menjadi data latih. Pada penelitian ini *cross validation* akan dibagi menjadi 10 *fold cross validation*. Dari 1040 data yang ada maka 1 bagian data akan terdiri dari 104 data dengan 10 bagian. Kemudian dilakukan pengukuran evaluasi untuk menghitung tingkat akurasi kebenaran proses klasifikasi digunakan data *confusion matrix*. *Confusion matrix* merupakan *output* dari klasifikasi SVM.