

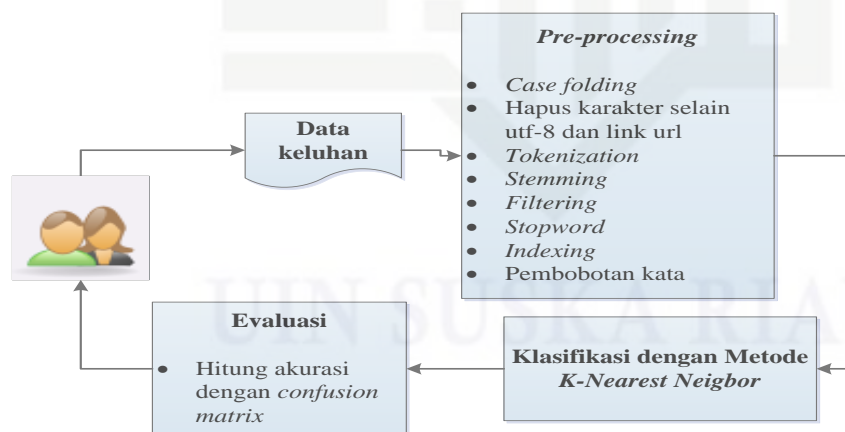
BAB IV

ANALISA DAN PERANCANGAN

Tahap ini merupakan tahap dimana dilakukan analisa dan perancangan terhadap sistem yang akan dibuat. Analisa merupakan kegiatan dalam penentuan tindakan atau keputusan yang akan menentukan dari rancangan sistem. Sementara tahap perancangan merupakan tahap kegiatan menentukan rincian sistem yang akan dibuat.

4.1 Analisa Sistem Klasifikasi Keluhan

Sistem klasifikasi keluhan mahasiswa ini memiliki tiga proses utama yaitu *pre-processing query*, klasifikasi dengan *K-Nearest Neighbor* dan evaluasi. Pada tahap *pre-processing* terdiri dari beberapa tahapan, yaitu *case folding*, menghapus karakter UTF-8, link URL, *tokenizing*, *stemming*, *filtering*, *stopword*, *indexing*, dan pembobotan kata, pada dokumen keluhan mahasiswa yang akan digunakan pada tahap klasifikasi. Tahap klasifikasi dengan metode *K-Nearest Neighbor* terdiri dari beberapa tahapan, yaitu identifikasi setiap kata dan kalimat pada *query* keluhan. Selanjutnya, pada tahap evaluasi akan diuji tingkat keakuratan dari hasil klasifikasi yang telah dilakukan. Gambar 4.1 menggambarkan alur kerja sistem secara keseluruhan.



Gambar 4.1 Alur Kerja Sistem Klasifikasi Keluhan

Sistem menerima inputan berupa *query* keluhan dan kemudian dilakukan tahapan *pre-processing* terhadap *query* tersebut. Selanjutnya, sistem akan menghitung kesamaan antara *query* dengan data latih yang telah dibangun.

Kemudian dilakukan identifikasi terhadap keluhan yaitu keluhan *krs*, keluhan *login*, dan keluhan nilai. Selanjutnya, cari jumlah terbanyak dari ketetangaan hasil indentifikasi keluhan. Pencarian tetangga terbanyak akan menentukan klasifikasi *query* keluhan tersebut. Hasil klasifikasi ini nantinya akan dievaluasi, untuk menentukan tingkat keakuratan klasifikasi dengan metode *K-Nearest Neighbor* dalam mengklasifikan keluhan mahasiswa. Untuk lebih jelasnya, setiap langkah dari analisa sistem klasifikasi keluhan mahasiswa ini akan dijelaskan dalam sub-bab di bawah ini.

4.1.1 Analisa Kebutuhan Data

Sebelum melakukan proses klasifikasi, hal pertama yang dilakukan pada penelitian ini adalah mengumpulkan data keluhan yang nantinya akan digunakan sebagai dataset. Data dikumpulkan secara manual melalui grup *facebook iraise helpdesk*. Data keluhan yang dikumpulkan berjumlah 190 data, waktu pengumpulan data keluhan antara bulan Februari sampai dengan Desember dan di simpan dengan format **xlsx*. Selanjutnya menentukan kategori keluhan yang akan dijadikan sebagai kelas pada proses klasifikasi menggunakan metode yang dibahas dalam penelitian ini. Adapun kategori dan ciri-ciri keluhan adalah sebagai berikut:

1. Keluhan KRS

Keluhan KRS ditandai dengan informasi yang ada dalam keluhan seperti berikut:

- a. Kesalahan terhadap kartu rencana studi
- b. Kesalahan terhadap KRS
- c. Kesalahan ambil matakuliah atau hapus matakuliah
- d. Tidak dapat mengambil matakuliah
- e. Jadwal bentrok atau jadwal pada KRS berubah
- f. Perubahan dosen pembimbing akademik
- g. Kesalahan persetujuan KRS oleh pembimbing akademik

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.

2. Keluhan *login*

Keluhan login ditandai dengan informasi yang ada dalam keluhan seperti berikut:

- a. Kesalahan *password* atau kesalahan *user id*
- b. Kesalahan lupa *password*
- c. Kesalahan *email* atau lupa *password email*
- d. Pembayaran spp, kecuali ketika login muncul anda belum melakukan pembayaran spp.

3. Keluhan nilai

Keluhan nilai ditandai dengan informasi yang ada dalam keluhan seperti berikut:

- a. Kesalahan nilai
- b. Kesalahan pada kartu hasil studi
- c. Kesalahan terhadap transkrip nilai

Penentuan kategori keluhan dipandu oleh Salti Reliani Primasta yang dilakukan pada tahun 2017, dengan mengisi *sample* data sebanyak 25 data keluhan. Berikut adalah contoh keluhan yang dan sudah di berikan kategori dapat dilihat pada Tabel 4.1 di bawah ini:

Tabel 4.1 Contoh Keluhan Berlabel

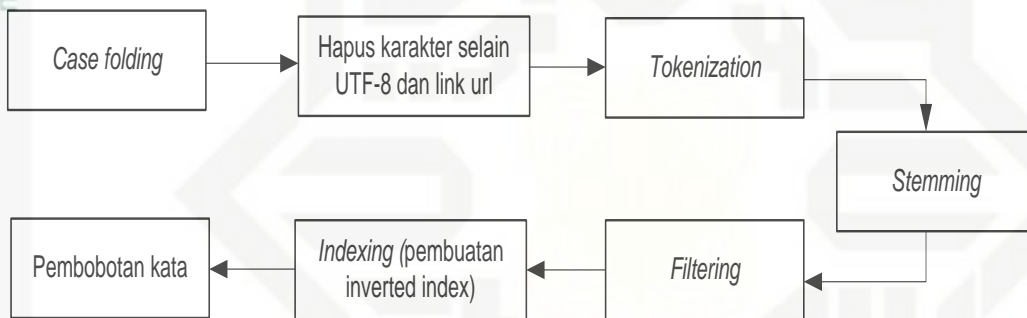
Dok(n)	Keluhan	Kategori
Dok(1)	Mau bilang, kalo iraise saya gak mau login, padahal passwordnya gak ada saya ganti, masih tetap nim password nya tu....mohon bantuannya agar bisa login	Login
Dok(2)	mohon bantuannya bapak/ibu krs saya tidak bisa login, karena password nya salah, padahal saya tidak ada mengganti password iraise saya, dan ketika saya reset password email yang tertera bukan email yang saya isi di profil iraise saya. mohon bantuan dan solusinya bapak/ibu	Login
Dok(3)	Saya daftar online knk tetapi tidak bisa karena krs tidak mencukupi padahal krs saya sudah melebihi syarat, Bagaimana ya solusi nya pak/ibu???	Krs
Dok(4)	Saya ada perbaikan krs dan sudah mnemui PA untuk persetujuan. Tapi belum di setujui. Apakah bisa saya langsung meminta proses oleh pihak admin.?	Krs
Dok(5)	nilai saya di iraise belum di input, padahal saya sudah mengajukan persyaratan nilai yang belum di input ke iraise, saya mengajukan persyaratannya sudah hampir satu bulan, mohon bantuannya kak...makasih..	Nilai
Dok(6)	Nilai saya satu belum keluar di khs online pak. Sudah dri 10 hari yang lalu saya urus cuma sampai skrng di sistem iraise	Nilai

Dok(n)	Keluhan	Kategori
	masih kosong atau nol. Gmna itu pak?	

Pada Tabel 4.1 adalah contoh data keluhan yang telah dikumpulkan dan sudah diberi label secara manual berdasarkan kategori keluhan yang didapat dari kuisisioner. Setelah data diberi label secara manual, selanjutnya adalah melakukan proses *pre-processing* terhadap keluhan.

4.1.2 Pre-Processing

Pre-processing merupakan langkah penting dalam melakukan klasifikasi yang bertujuan untuk menyamakan data, dan membersihkan data dari unsur-unsur yang tidak di butuhkan. Gambar 4.2 merupakan langkah secara umum tahapan *text pre-processing*.



Gambar 4.2 Proses Text Pre-Processing

Berikut ini penjelasan gambar 4.2 yang merupakan tahapan *text pre-processing* :

1. Case folding

Proses *Case folding* adalah proses penyeragaman bentuk huruf dengan mengubah semua huruf kapital menjadi huruf kecil, karena tidak semua data keluhan konsisten dalam penggunaan huruf. Oleh karena itu peran *case folding* dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (biasanya huruf kecil). Dari contoh keluhan di atas berikut adalah hasil proses *case folding* seperti pada Tabel 4.2 di bawah ini.

Tabel 4.2 Hasil Proses Case Folding

Dok(n)	Keluhan	Kategori
Dok(1)	mau bilang, kalo iraise saya gak mau login,padahal passwordnya gak ada saya ganti, masih tetap nim passwordnya tu....mohon bantuannya agar bisa login	login

Dok(n)	Keluhan	Kategori
Dok(2)	mohon bantuannya bapak/ibu krs saya tidak bisa login, karena password nya salah, padahal saya tidak ada mengganti password iraise saya, dan ketika saya reset password email yang tertera bukan email yang saya isi di profil iraise saya.mohon bantuan dan solusinya bapak/ibu	login
Dok(3)	saya daftar online kkn tetapi tidak bisa karena krs tidak mencukupi padahal krs saya sudah melebihi syarat,bagaimana ya solusi nya pak/ibu???	Krs
Dok(4)	saya ada perbaikan krs dan sudah menemui pa untuk persetujuan. tapi blum di setujui. apakah bisa saya langsung meminta proses oleh pihak admin.?	Krs
Dok(5)	nilai saya di iraise belum di input,padahal saya udah mengajukan persyaratan nilai yang belum di input ke iraise, saya mengajukan persyaratannya udah hampir satu bulan, mohon bantuannya kak...makasih..	Nilai
Dok(6)	nilai saya satu belum keluar di khs online pak. udah dri 10 hari yang lalu saya urus cuma sampai skrng di sistem iraise masih kosong atau nol. gmna itu pak?	Nilai

2. Menghapus karakter UTF-8 dan *link url*

Proses ini akan menghapus karakter yang tidak relevan yang dianggap dapat mempengaruhi akurasi dari sistem. Hasil proses tahapan ini seperti pada Tabel 4.3 di bawah ini.

Tabel 4.3 Proses Menghapus Karakter UTF-8 dan *link url*

Dok(n)	Keluhan	Kategori
Dok(1)	mau bilang kalo iraise saya gak mau login padahal passwordnya gak ada saya ganti masih tetap nim password nya tu mohon bantuannya agar bisa login	login
Dok(2)	mohon bantuannya bapak ibu krs saya tidak bisa login karena password nya salah padahal saya tidak ada mengganti password iraise saya dan ketika saya reset password email yang tertera bukan email yang saya isi di profil iraise saya mohon bantuan dan solusinya bapak ibu	login
Dok(3)	saya daftar online kkn tetapi tidak bisa karena krs tidak mencukupi padahal krs saya sudah melebihi syarat bagaimana ya solusi nya pak ibu	Krs
Dok(4)	saya ada perbaikan krs dan sudah menemui pa untuk persetujuan tapi blum di setujui apakah bisa saya langsung meminta proses oleh pihak admin	Krs
Dok(5)	nilai saya di iraise belum di input padahal saya udah mengajukan persyaratan nilai yang belum di input ke iraise saya mengajukan persyaratannya udah hampir satu bulan mohon bantuannya kak makasih	Nilai
Dok(6)	nilai saya satu belum keluar di khs online pak udah dri hari yang lalu saya urus cuma sampai skrng di sistem iraise masih kosong atau nol gmna itu pak	Nilai

3. Tokenization

Proses *tokenization* adalah proses memisahkan deretan kata menjadi potongan kata atau token. Pada tahap ini seluruh kalimat pada dokumen akan dipisahkan menjadi token atau potongan kata tunggal. Hasil proses *tokenization* seperti pada Tabel 4.4 di bawah ini.

Table 4.4 Proses Tokenization

Dok(1)	Dok(2)	Dok(3)	Dok(4)	Dok(5)	Dok(6)
mau	mohon	saya	saya	nilai	nilai
bilang	bantuannya	daftar	ada	saya	saya
kalo	bapak	online	Perbaikan	di	satu
iraise	ibu	kkn	krs	iraise	belum
saya	krs	tetapi	dan	belum	keluar
gak	saya	tidak	sudah	di	di
mau	tidak	bisa	menemui	input	khs
login	bisa	karena	pa	padahal	online
padahal	login	krs	untuk	saya	pak
passwordnya	karena	tidak	persetujuan	sudah	sudah
gak	password	mencukupi	tapi	mengajukan	dri
ada	nya	padahal	belum	persyaratan	hari
saya	salah	krs	di	nilai	yang
ganti	padahal	saya	setujui	yang	lalu
masih	saya	sudah	apakah	belum	saya
tetap	tidak	melebihi	bisa	di	urus
nim	ada	syarat	saya	input	cuma
password	mengganti	bagaimana	langsung	ke	sampai
nya	password	ya	meminta	iraise	skrg
tu	iraise	solusi	proses	saya	di
mohon	saya	nya	oleh	mengajukan	sistem
bantuannya	dan	pak	pihak	persyaratannya	iraise
agar	ketika	ibu	admin	sudah	masih
bisa	saya			hampir	kosong
login	reset			satu	atau
	password			bulan	nol
	email			mohon	gmna
	yang			bantuannya	itu
	tertera			kak	pak
	bukan			makasih	
	email				
	yang				
	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.

Dok(1)	Dok(2)	Dok(3)	Dok(4)	Dok(5)	Dok(6)
	isi				
	di				
	profil				
	iraise				
	saya				
	mohon				
	bantuan				
	dan				
	solusinya				
	bapak				
	ibu				

4. Stemming

Proses *stemming* merupakan proses mengubah kata-kata yang mengandung imbuhan, awalan (prefix), akhiran (suffix), sisipan (infix), dan awalan akhiran (konfix) menjadi kata dasar dengan menggunakan algoritma *stemming* Nazief dan Adriani. Berdasarkan contoh keluhan di atas, maka hasil *stemming*-nya dapat dilihat pada Tabel 4.5 di bawah ini.

Tabel 4.5 Hasil Proses Stemming

Dok(1)	Dok(2)	Dok(3)	Dok(4)	Dok(5)	Dok(6)
mau	mohon	saya	saya	nilai	nilai
bilang	bantu	daftar	ada	saya	saya
kalo	bapak	online	baik	di	satu
iraise	ibu	kkn	krs	iraise	belum
saya	krs	tetapi	dan	belum	keluar
gak	saya	tidak	sudah	di	di
mau	tidak	bisa	temu	input	khs
login	bisa	karena	pa	padahal	online
padahal	login	krs	untuk	saya	pak
passwordnya	karena	tidak	setuju	sudah	sudah
gak	password	cukup	tapi	ajuk	dri
ada	nya	padahal	belum	syarat	hari
saya	salah	krs	di	nilai	yang
ganti	padahal	saya	setuju	yang	lalu
masih	saya	sudah	apakah	belum	saya
tetap	tidak	lebih	bisa	di	urus
nim	ada	syarat	saya	input	cuma
password	ganti	bagaimana	langsung	ke	sampai
nya	password	ya	minta	iraise	skrg

- 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.

Dok(1)	Dok(2)	Dok(3)	Dok(4)	Dok(5)	Dok(6)
tu	iraise	solusi	proses	saya	di
mohon	saya	nya	oleh	ajuk	sistem
bantu	dan	pak	pihak	syarat	iraise
agar	ketika	ibu	admin	sudah	masih
bisa	saya			hampir	kosong
login	reset			satu	atau
	password			bulan	nol
	email			mohon	gmna
	yang			bantu	itu
	tera			kak	pak
	bukan			makasih	
	email				
	yang				
	saya				
	isi				
	di				
	profil				
	iraise				
	saya				
	mohon				
	bantu				
	dan				
	solusi				
	bapak				
	ibu				

5. Filtering

Tahap *filtering* yang dilakukan pada penelitian ini adalah menghilangkan kata-kata yang sering muncul atau kata-kata yang tidak penting yang dapat menekan penggunaan waktu dalam menentukan hasil, tahapan ini juga disebut tahapan *stopword list*. *Stopword* berisi data kata-kata yang tidak begitu berpengaruh atau tidak memiliki arti yang penting pada dokumen. *Stopword* yang digunakan mengacu pada *stopword* bahasa Indonesia yang disediakan oleh link <https://code.google.com/archive/p/stop-words/>. Daftar kata *stopword* dapat dilihat pada Lampiran A. Hasil dari tahapan *filtering* pada analisa ini akan diperlihatkan oleh Tabel 4.6 di bawah ini.

Tabel 4.6 Hasil Proses *Filtering* dan Menghilangkan *Stopword*

Dok(1)	Dok(2)	Dok(3)	Dok(4)	Dok(5)	Dok(6)
bilang	mohon	daftar	baik	nilai	nilai
kalo	bantu	online	krs	iraise	satu
iraise	bapak	kkn	temu	input	keluar
gak	ibu	krs	pa	ajuk	khs
login	krs	cukup	setuju	syarat	online
passwordnya	login	krs	setuju	nilai	pak
gak	password	syarat	langsung	input	dri
ganti	nya	ya	minta	iraise	hari
tetap	salah	solusi	proses	ajuk	urus
nim	ganti	nya	pihak	syarat	skrg
password	password	pak	admin	satu	sistem
nya	iraise	ibu		bulan	iraise
tu	reset			mohon	kosong
mohon	password			bantu	nol
bantu	email			kak	gmna
login	tera			makasih	pak
	email				
	isi				
	profil				
	iraise				
	mohon				
	bantu				
	solusi				
	bapak				
	ibu				

6. *Indexing* (Pembuatan *Inverted Index*)

Pada tahap ini, dilakukan proses pengindeksan pada setiap token. Pada *inverted index* akan tersimpan informasi berupa kata, id dokumen yang mengandung kata tersebut dan frekuensi kemunculan kata atau token tersebut dalam suatu dokumen. Hasil proses *Indexing* dapat dilihat pada Tabel 4.7 berikut.

Table 4.7 Hasil Proses *Indexing* (Pembuatan *Inverted Index*)

Kosa Kata	(dokumen : frekuensi)
bilang	1:1
kalo	1:1
iraise	1:1 2:2 5:1
gak	1:2

Kosa Kata	(dokumen : frekuensi)		
login	1:2	2:1	
ganti	1:1	2:1	
tetap	1:1		
nim	1:1		
password	1:2	2:3	
mohon	1:1	2:2	5:1
bantu	1:1	2:2	5:1
bapak	2:2		
ibu	2:2	3:1	
krs	2:1	3:2	4:1
salah	2:1		
reset	2:1		
email	2:2		
tera	2:1		
isi	2:1		
profil	3:1		
solusi	2:1	3:1	
daftar	3:1		
online	3:1		6:1
kkn	3:1		
cukup	3:1		
syarat	3:1	5:3	
ya	3:1		
pak	3:1		6:2
baik	4:1		
temu	4:1		
pa	4:1		
setuju	4:2		
langsung	4:1		
minta	4:1		
proses	4:1		
pihak	4:1		
admin	4:1		
nilai	5:2	6:1	
input	5:2		
ajuk	5:2		
satu	5:1	6:2	
bulan	5:1		
kak	5:1		
makasih	5:1		
keluar	6:1		

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.

Kosa Kata	(dokumen : frekuensi)
khs	6:1
online	6:1
dri	6:1
hari	6:1
urus	6:1
skrg	6:1
sistem	6:1
kosong	6:1
nol	6:1
gmna	6:1

7. Pembobotan kata

Hasil dari *indexing* pada proses sebelumnya akan dihitung nilai bobotnya. Pada penelitian ini digunakan dua jenis perhitungan pembobotan. yakni pembobotan lokal dengan menggunakan *tf* (*term frequency*) dan pembobotan global dengan menggunakan *idf* (*inverse dokumen frequency*) menggunakan Persamaan (2.1). Hasil pembobotan kata dapat dilihat pada Tabel 4.8 berikut.

Table 4.8 Hasil Pembobotan Kata

Kosa Kata	<i>tf</i> (D1)	<i>tf</i> (D2)	<i>tf</i> (D3)	<i>tf</i> (D4)	<i>tf</i> (D5)	<i>tf</i> (D6)	<i>Df</i>
bilang	1	0	0	0	0	0	1
kalo	1	0	0	0	0	0	1
iraise	1	2	0	0	1	0	3
gak	2	0	0	0	0	0	1
login	2	1	0	0	0	0	2
ganti	1	1	0	0	0	0	2
tetap	1	0	0	0	0	0	1
nim	1	0	0	0	0	0	1
password	2	3	0	0	0	0	2
mohon	1	2	0	0	1	0	3
bantu	1	2	0	0	1	0	3
agar	1	0	0	0	0	0	1
bapak	0	2	0	0	0	0	1
ibu	0	2	1	0	0	0	2
krs	0	1	2	1	0	0	3
salah	0	1	0	0	0	0	1
reset	0	1	0	0	0	0	1
email	0	2	0	0	0	0	1
tera	0	1	0	0	0	0	1

Kosa Kata	$tf(D1)$	$tf(D2)$	$tf(D3)$	$tf(D4)$	$tf(D5)$	$tf(D6)$	Df
isi	0	1	0	0	0	0	1
profil	0	0	1	0	0	0	1
solusi	0	1	1	0	0	0	2
daftar	0	0	1	0	0	0	1
online	0	0	1	0	0	1	2
kkn	0	0	1	0	0	0	1
cukup	0	0	1	0	0	0	1
lebih	0	0	1	0	0	0	1
syarat	0	0	1	0	3	0	2
ya	0	0	1	0	0	0	1
pak	0	0	1	0	0	2	2
baik	0	0	0	1	0	0	1
temu	0	0	0	1	0	0	1
pa	0	0	0	1	0	0	1
setuju	0	0	0	2	0	0	1
langsung	0	0	0	1	0	0	1
minta	0	0	0	1	0	0	1
proses	0	0	0	1	0	0	1
pihak	0	0	0	1	0	0	1
admin	0	0	0	1	0	0	1
nilai	0	0	0	0	2	1	2
input	0	0	0	0	2	0	1
ajuk	0	0	0	0	2	0	1
satu	0	0	0	0	1	2	2
bulan	0	0	0	0	1	0	1
kak	0	0	0	0	1	0	1
makasih	0	0	0	0	1	0	1
keluar	0	0	0	0	0	1	1
khs	0	0	0	0	0	1	1
online	0	0	0	0	0	1	1
dri	0	0	0	0	0	1	1
hari	0	0	0	0	0	1	1
urus	0	0	0	0	0	1	1
skrg	0	0	0	0	0	1	1
sistem	0	0	0	0	0	1	1
kosong	0	0	0	0	0	1	1
nol	0	0	0	0	0	1	1
gmna	0	0	0	0	0	1	1

Hak Cipta Dilindungi Undang-Undang

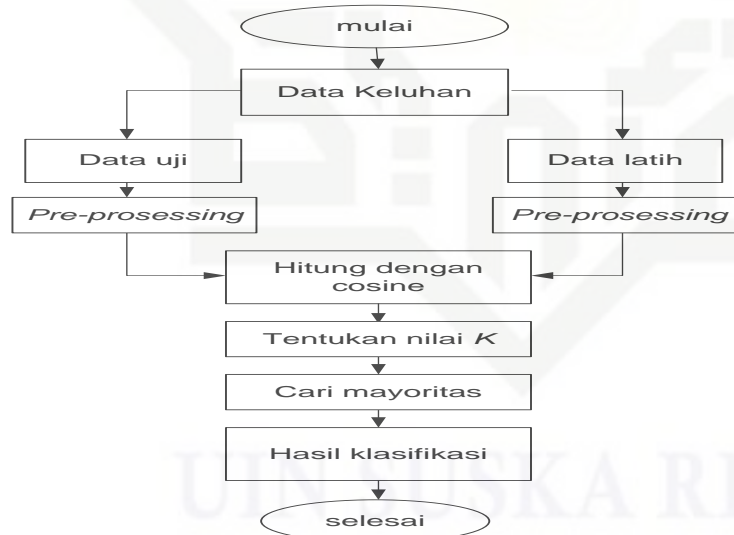
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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

Keterangan :

- Kosa kata : kata tunggal yang telah terpisah dari dokumen keluhan
 Df : *Document frequency* atau jumlah dokumen dalam koleksi dokumen yang mengandung kosa kata
 idf : *inverse document frequency*
 $tf(n)$: *Term frequency* (frekuensi kata) pada dokumen keluhan ke- n

4.1.3 Klasifikasi Dengan Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* adalah algoritma untuk mengklasifikasi objek baru berdasarkan atribut dan *training samples* (data latih). Dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada *K-Nearest Neighbor*. Setelah dokumen melalui tahapan proses pada proses sebelumnya, maka pada bagian ini akan dijelaskan bagaimana penggunaan metode *K-Nearest Neighbor*. Untuk lebih jelas mengenai cara kerja algoritma *K-Nearest Neighbor* ini akan dijelaskan pada *flowchart* yang diperlihatkan pada Gambar 4.3 di bawah ini.



Gambar 4.3 Flowchart Alur Klasifikasi *K - Nearest Neighbor*

Pada proses klasifikasi dengan algoritma *K - Nearest Neighbor* dibagi menjadi 4 proses utama yaitu:

1. Pembagian data

Pada tahap ini, seluruh data yang telah didapat akan dibagi menjadi data latih dan data uji.

2. Melakukan *text pre-processing* pada data latih dan data uji
3. Menghitung ketetangaan terhadap latih yang telah dibangun. Setiap dokumen akan di hitung jaraknya dengan query menggunakan persamaan 2.4 *Cosine Similarity*
4. Menentukan nilai *k* pada algoritma *K-Nearest Neighbor*. Nilai *k* merupakan parameter yang akan membatasi ketetangaan terdekat yang akan diproses pada tahap selanjutnya.
5. Menyimpulkan klasifikasi keluhan dengan persamaan 2.5.

Untuk lebih jelasnya, akan dijelaskan tahapan Klasifikasi dengan algoritma *K-Nearest Neighbor* sebagai lanjutan contoh dari proses *Pre-processing*. Sebagai contoh *query* data uji yang di gunakan adalah sebagai berikut:

Query Data Uji : kok iraise saya tidak bisa login ,padahal udah di reset password nya gak bisa juga login

- a. Menghitung ketetangaan terhadap latih yang telah dibangun.

Sebelum menghitung *Cosine Similarity* terlebih dahulu dihitung bobot kata setiap dokumen. Berikut ini hasil bobot kata per dokumen pada Tabel 4.9

Table 4.9 Hasil Bobot Kata terhadap Kata Hasil Indexing

Kosa Kata	tf							Df	Log n/df	Wdt=Tf.idf						
	q	1	2	3	4	5	6			W (Q)	W (D1)	W (D2)	W (D3)	W (D4)	W (D5)	W (D6)
bilang		1						1	0,845	0	0,845	0	0	0	0	0
kalo		1						1	0,845	0	0,845	0	0	0	0	0
iraise	1	1	2					3	0,367	0,367	0,367	0,734	0	0	0	0
gak	1	2						2	0,544	0,544	1,088	0	0	0	0	0
login	2	2	1					3	0,367	0,734	0,734	0,367	0	0	0	0
ganti		1	1					2	0,544	0	0,544	0,544	0	0	0	0
tetap		1						1	0,845	0	0,845	0	0	0	0	0
nim		1						1	0,845	0	0,845	0	0	0	0	0
password	1	2	3					3	0,301	0,301	0,301	0,903	0	0	0	0
mohon		2	1			1		3	0,367	0	0,734	0,367	0	0	0,367	0
bantu		1	2			1		3	0,367	0	0,367	0,734	0	0	0,367	0
bapak			2					1	0,845	0	0	1,690	0	0	0	0
ibu			2	1				2	0,544	0	0	1,088	0,544	0	0	0
krs			1	2	1			3	0,367	0	0	0,367	0,734	0,367	0	0
salah			1					1	0,845	0	0	0,845	0	0	0	0
reset	1	1						2	0,544	0,544	0	0,544	0	0	0	0
email			2					1	0,845	0	0	1,690	0	0	0	0

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

Kosa Kata	tf								Idf	Wdt=Tf.idf						
	q	1	2	3	4	5	6	Df		Log n/df	W (Q)	W (D1)	W (D2)	W (D3)	W (D4)	W (D5)
tera			1					1	0,845	0	0	0,845	0	0	0	0
isi			1					1	0,845	0	0	0,845	0	0	0	0
profil				1				1	0,845	0	0	0	0,845	0	0	0
solusi			1	1				2	0,544	0	0	0,544	0,544	0	0	0
daftar				1				1	0,845	0	0	0	0,845	0	0	0
online				1			1	2	0,544	0	0	0	0,544	0	0	0,544
kkn				1				1	0,845	0	0	0	0,845	0	0	0
cukup				1				1	0,845	0	0	0	0,845	0	0	0
syarat				1		3		2	0,544	0	0	0	0,544	0	1,632	0
ya				1				1	0,845	0	0	0	0,845	0	0	0
pak				1			2	2	0,544	0	0	0	0,544	0	0	1,088
baik					1			1	0,845	0	0	0	0	0,845	0	0
temu					1			1	0,845	0	0	0	0	0,845	0	0
pa					1			1	0,845	0	0	0	0	0,845	0	0
setuju					1			1	0,845	0	0	0	0	0,845	0	0
langsung					1			1	0,845	0	0	0	0	0,845	0	0
minta					1			1	0,845	0	0	0	0	0,845	0	0
proses					1			1	0,845	0	0	0	0	0,845	0	0
pihak					1			1	0,845	0	0	0	0	0,845	0	0
admin					1			1	0,845	0	0	0	0	0,845	0	0
nilai						2	1	2	0,544	0	0	0	0	0	1,088	0,544
input						2		1	0,845	0	0	0	0	0	1,690	0
ajuk						2		1	0,845	0	0	0	0	0	1,690	0
satu						1	2	2	0,544	0	0	0	0	0	0,544	1,088
bulan						1		1	0,845	0	0	0	0	0	0,845	0
kak						1		1	0,845	0	0	0	0	0	0,845	0
makasih						1		1	0,845	0	0	0	0	0	0,845	0
keluar							1	1	0,845	0	0	0	0	0	0	0,845
khs							1	1	0,845	0	0	0	0	0	0	0,845
online							1	1	0,845	0	0	0	0	0	0	0,845
dri							1	1	0,845	0	0	0	0	0	0	0,845
hari							1	1	0,845	0	0	0	0	0	0	0,845
urus							1	1	0,778	0	0	0	0	0	0	0,778
skrg							1	1	0,778	0	0	0	0	0	0	0,778
sistem							1	1	0,778	0	0	0	0	0	0	0,778
kosong							1	1	0,778	0	0	0	0	0	0	0,778
nol							1	1	0,778	0	0	0	0	0	0	0,778
gmna							1	1	0,778	0	0	0	0	0	0	0,778

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.

Selanjutnya berdasarkan perhitungan bobot tersebut, hitung panjang vector setiap dokumen. Tahapan perhitungan ketetanggan *query* data uji terhadap data latih dengan persamaan (2.4) yaitu *Cosine Similarity*. Dapat dijelaskan dengan uraian sebagai berikut :

1. Perhitungan kemiripan query dengan dokumen (1)

$$R(Q, d1) = \frac{(0,367x1) + (0,544x2) + (0,734x2) + (0,301x1) + (0,243x1) + (0,554x1)}{\sqrt{7,974} \times \sqrt{1,414}}$$

$$R(Q, d1) = \frac{0,367 + 1,088 + 1,468 + 0,301 + 0,301 + 0,544}{2,823 \times 1,189}$$

$$R(Q, d1) = \frac{4,011}{3,356}$$

$$R(Q, d1) = 1,195$$

2. Perhitungan kemiripan query dengan dokumen (2)

$$R(Q, d2) = \frac{(0,367x2) + (0,544x0) + (0,734x1) + (0,301x3) + (0,243x1) + (0,554x0)}{\sqrt{12,281} \times \sqrt{1,414}}$$

$$R(Q, d2) = \frac{0,734 + 0 + 0,734 + 0,903 + 0,243 + 0}{3,504 \times 1,189}$$

$$R(Q, d2) = \frac{2,614}{4,166}$$

$$R(Q, d2) = 0,627$$

3. Perhitungan kemiripan query dengan dokumen (3)

$$R(Q, d3) = \frac{(0,367x0) + (0,544x0) + (0,734x0) + (0,301x0) + (0,243x1) + (0,554x0)}{\sqrt{6,361} \times \sqrt{1,414}}$$

$$R(Q, d3) = \frac{0 + 0 + 0 + 0 + 0,243 + 0}{2,522 \times 1,189}$$

$$R(Q, d3) = \frac{0,243}{2,998}$$

$$R(Q, d3) = 0,081$$

4. Perhitungan kemiripan query dengan dokumen (4)

$$R(Q, d4) = \frac{(0,367x0) + (0,544x0) + (0,734x0) + (0,301x0) + (0,243x0) + (0,554x0)}{\sqrt{7,274} \times \sqrt{1,414}}$$

$$R(Q, d4) = \frac{0 + 0 + 0 + 0 + 0 + 0}{2,697 \times 1,189}$$

$$R(Q, d4) = \frac{0}{3,206}$$

$$R(Q, d4) = 0$$

5. Perhitungan kemiripan query dengan dokumen (5)

$$R(Q, d5)$$

$$= \frac{(0,367x0) + (0,544x0) + (0,734x0) + (0,301x0) + (0,243x0) + (0,554x0)}{\sqrt{12,266} \times \sqrt{1,414}}$$

$$R(Q, d5) = \frac{0 + 0 + 0 + 0 + 0 + 0}{3,502 \times 1,189}$$

$$R(Q, d5) = \frac{0}{4,163}$$

$$R(Q, d5) = 0$$

6. Perhitungan kemiripan query dengan dokumen (6)

$$R(Q, d6)$$

$$= \frac{(0,367x0) + (0,544x0) + (0,734x0) + (0,301x0) + (0,243x0) + (0,554x0)}{\sqrt{10,161} \times \sqrt{1,414}}$$

$$R(Q, d6) = \frac{0 + 0 + 0 + 0 + 0 + 0}{3,187 \times 1,189}$$

$$R(Q, d6) = \frac{0}{3,789}$$

$$R(Q, d6) = 0$$

Hasil dari kemiripan *query* dengan dokumen data latih menunjukkan hasil sebagai berikut :

$$D_1 = 1,195, D_2 = 0,627, D_3 = 0,081, D_4 = 0, D_5 = 0, D_6 = 0$$

b. Menentukan nilai *k* pada algoritma *K-Nearest Neighbor*. Nilai *k* merupakan parameter yang akan membatasi ketetanggaan terdekat yang akan diproses pada tahapan selanjutnya.

Ditentukan nilai *k* pada algoritma *K-Nearest Neighbor* adalah 3. Sehingga berdasarkan nilai tersebut, maka ketetanggaan terdekat dengan *query* data uji adalah sebagai berikut:

$$D_1 = 1,195, D_2 = 0,627, D_3 = 0,081$$

c. Mengurutkan jarak terkecil dari data latih sehingga setelah di urutkan menjadi seperti berikut:

$$D_3 = 0,081, D_2 = 0,627, D_1 = 1,195$$

d. Cari jumlah terbanyak dari tetangga dengan persamaan (2.5). Kemudian tetapkan kategori.

$$D_3 = \text{Krs}, D_2 = \text{Login}, D_1 = \text{Login}$$

$$D_{nn} (C_1, C_2) = \text{Krs} = 1, \text{Login} = 2, \text{Nilai} = 0$$

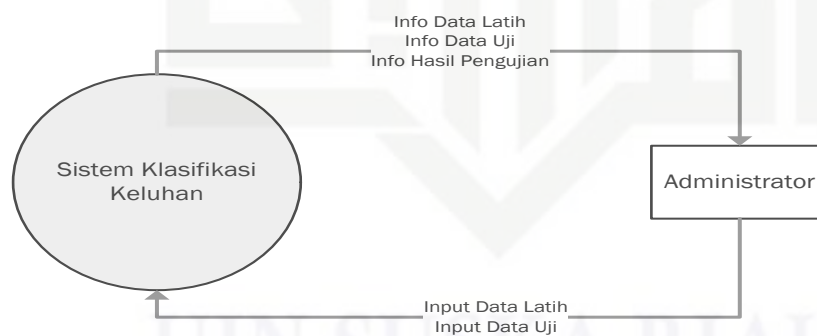
Sehingga dapat disimpulkan bahwa *query* data uji memiliki kategori **Login**.

4.1.4 Analisa Fungsional Sistem

Analisa fungsional sistem akan menjelaskan mengenai perancangan sistem yang akan dibuat menggunakan *Flow Diagram* (FD) meliputi *Context Diagram*, *Data Flow Diagram* (DFD).

a. *Context Diagram*

Context Diagram menggambarkan aliran fungsional dalam sebuah proses pada sistem atau aplikasi. Berikut akan dijelaskan semua proses yang terjadi pada sistem. *Context Diagram* akan dijelaskan seperti pada Gambar 4.4 di bawah ini.



Gambar 4.4 Context Diagram

Entitas yang berinteraksi dengan sistem ini adalah *Administrator*. Dalam sistem ini, *administrator* dapat melakukan penghapusan data.

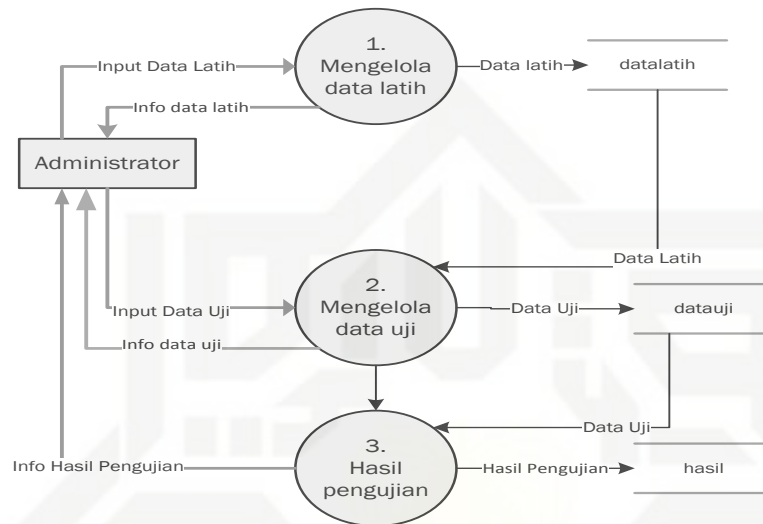
b. *Data Flow Diagram* (DFD)

Data Flow Diagram (DFD) sering digunakan untuk menggambarkan suatu sistem yang telah ada atau sistem baru yang akan dikembangkan secara logika

tanpa mempertimbangkan lingkungan fisik dimana data tersebut mengalir, atau lingkungan fisik dimana data tersebut tersimpan.

1. Data Flow Diagram (DFD) Level 1

Gambar 4.5 di bawah ini adalah gambaran DFD *level 1* dari sistem klasifikasi keluhan Mahasiswa.



Gambar 4.5 DFD Level 1

Pada Tabel 4.10 di bawah ini adalah proses DFD *Level 1*. Terdapat 3 (tiga) proses yaitu mengelola data latihan, mengelola data uji dan hasil pengujian.

Tabel 4.10 Proses DFD Level 1

No	Nama Proses	Deskripsi
1.	Data Latihan	Proses pengolahan data latihan
2.	Pengujian	Proses pengolahan data uji
3.	Hasil Pengujian	Tampilan hasil akurasi dalam <i>confusion matrix</i> .

Pada Tabel 4.11 di bawah ini adalah aliran data DFD *level 1*. Terdapat 5 (lima) aliran data yaitu data latihan, info data latihan, data uji, info data uji, dan hasil pengujian.

Tabel 4.11 Aliran Data DFD level 1

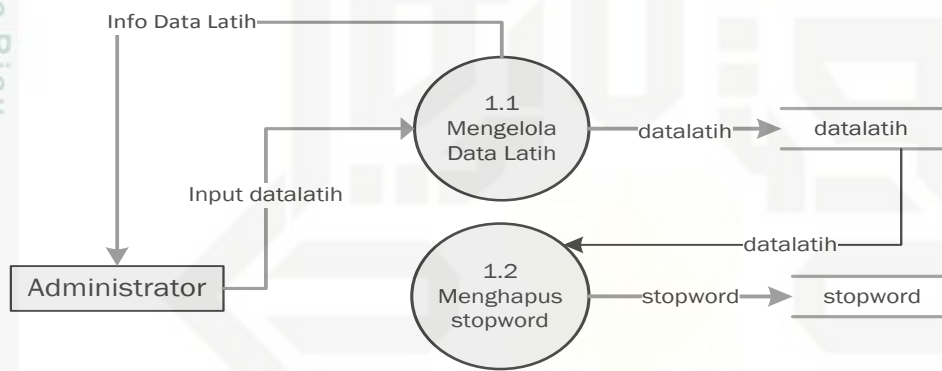
Nama Aliran Data	Deskripsi
Input Data Latihan	Data latihan yang diinputkan oleh admin

Nama Aliran Data	Deskripsi
Input Data Uji	Data uji yang diinputkan oleh admin
Info Hasil Pengujian	Info data hasil pengujian

2. Data Flow Diagram (DFD) Level 2 proses 1 (Data Latih)

Pada DFD level 2 proses 1 ini, terdiri dari proses Data latih dan *stopword*. *Datastore* terdiri dari *datalatih* dan *stopword*. Berikut ini adalah proses DFD level 2 dari proses data latih .

Pada Gambar 4.6 di bawah ini adalah gambar DFD Level 2 Proses 1 (Data Latih). Terdapat 2(dua) proses yaitu Data Latih dan *stopword*.



Gambar 4.6 DFD Level 2 Proses 1

Pada Tabel 4.12 di bawah ini adalah proses DFD Level 2 Proses 1. Terdapat 2 proses yaitu Data Latih dan *stopword*.

Tabel 4.12 Proses DFD Level 2 Proses 1

No	Nama Proses	Deskripsi
1.	Data Latih	Proses pengolahan data latih
2.	Stopword	Proses eliminasi kata data latih

Pada Tabel 4.13 di bawah ini adalah aliran data DFD level 2 Proses 2. Terdapat 2 (dua) proses yaitu data latih dan *stopword*.

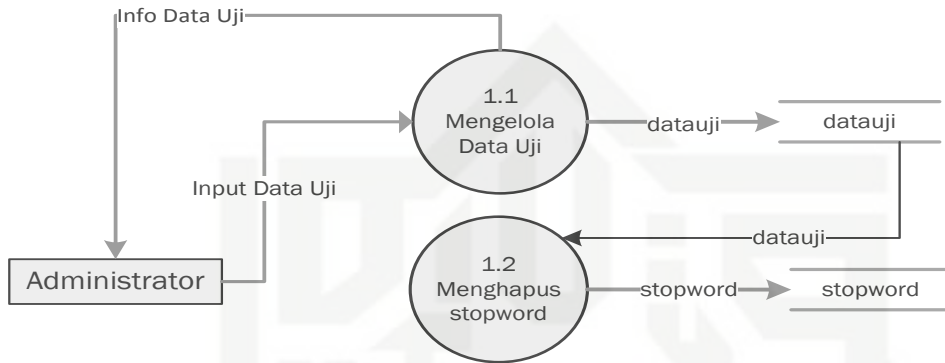
Tabel 4.13 Aliran Data DFD level 2 Proses 2

Nama Aliran Data	Deskripsi
Input Data Latih	Data latih yang diinputkan oleh admin
Info Data Latih	Info data latih

3. Data Flow Diagram (DFD) Level 2 proses 2 (Pengujian)

Pada DFD level 2 proses 2 ini, terdiri dari proses Data uji dan *stopword*. *Datastore* terdiri dari *datauji* dan *stopword*. Berikut ini adalah proses DFD level 2 dari proses pengujian.

Pada Gambar 4.16 di bawah ini adalah gambar DFD Level 2 Proses 2 (Data Uji). Terdapat 2 (dua) proses yaitu Data Uji dan *stopword*.



Gambar 4.7 DFD Level 2 Proses 2

Pada Tabel 4.14 di bawah ini proses DFD Level 2 Proses 2. Terdapat 2 (dua) proses yaitu Data Uji dan *stopword*.

Tabel 4.14 Proses DFD Level 2 Proses 2

No	Nama Proses	Deskripsi
1.	Data Uji	Proses pengolahan data uji
2.	Stopword	Proses eliminasi kata data uji

Pada Tabel 4.15 di bawah ini adalah Aliran Data DFD level 2 Proses 2. Terdapat 2 (tiga) aliran data yaitu data uji dan info data uji.

Tabel 4.15 Aliran Data DFD level 2 Proses 2

Nama Aliran Data	Deskripsi
Data Uji	Proses input Data Uji
Info data uji	Info tampilan data uji

4.2 Perancangan Sistem

Pada tahap ini akan dibahas tentang perancangan sistem klasifikasi keluhan mahasiswa berdasarkan tahapan analisa yang telah dilakukan

sebelumnya. Adapun perancangan yang akan dibuat adalah perancangan permodelan *flat file*, perancangan struktur menu dan perancangan *interface*.

4.2.1 Perancangan *Flat file* (File Teks)

Dalam penelitian ini, akan menggunakan *flat file* yang bertindak sebagai penyimpanan yang akan mendukung pemrosesan klasifikasi keluhan. Adapun penggunaan *flat file* yang akan digunakan adalah sebagai penyimpanan dokumen latih, dokumen data uji, hasil indexing, kumpulan kata *stopword* dan kamus bahasa Indonesia, Untuk lebih jelasnya dapat dilihat pada Tabel 4.16 berikut.

Tabel 4.16 Rancangan *Flat File* (File Teks)

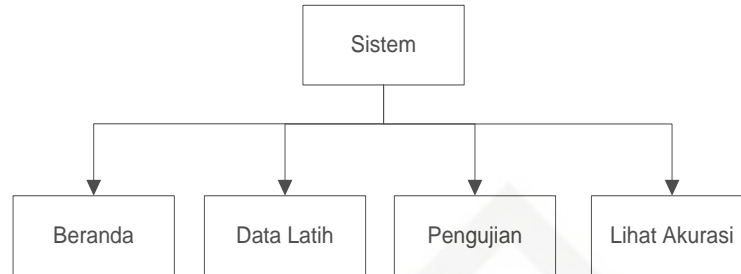
Nama File	Fungsi Penyimpanan
file-latih.txt	Sebagai penyimpanan data latih
file-uji.txt	Sebagai penyimpanan data uji
file-latih-index.txt	Sebagai penyimpanan hasil indexing
file-stopword.txt	Sebagai penyimpanan kata <i>stopword</i>
kamus-ind.txt	Sebagai penyimpanan kata dasar

Setiap *flat file* memiliki fungsi penyimpanan tersendiri dan memuat informasi yang berkaitan terhadap fungsi penyimpanan. Pada file-latih.txt akan memuat informasi id dokumen, kategori keluhan dan isi keluhan. Pada file-uji.txt akan memuat informasi id dokumen, kategori keluhan, isi keluhan dan hasil klasifikasi dengan algoritma *K-Nearest Neighbor*. Sedangkan file-latih-index.txt merupakan penyimpanan dari hasil indexing terhadap data latih dimana memuat informasi kata, id dokumen, *trem frequency (tf)*, *document frequency (df)*. Sementara itu file-stopword.txt dan kamus-ind.txt akan digunakan sebagai penyimpanan pada proses *pre-processing* yaitu daftar kata *stopword* pada tahapan *filtration* dan kamus indonesia sebagai penyimpanan kata dasar pada proses *steaming* yang masing-masing memuat informasi kata.

4.2.2 Perancangan Struktur Menu

Perancangan struktur menu menggambarkan hubungan antara halaman dengan halaman yang lainnya. Menu merupakan salah satu bagian penting dalam merancang sebuah *interface* pada sistem, karena melalui menu dapat dilihat bagaiman struktur sistem ini terbentuk.

Adapun struktur menu pada sistem klasifikasi ini terdiri dari Beranda, Data Latih, Pengujian dan Lihat Akurasi. Rancangan struktur menu dapat dilihat pada Gambar 4.8 berikut.



Gambar 4.8 Rancangan Struktur Menu

4.2.3 Perancangan *Interface*

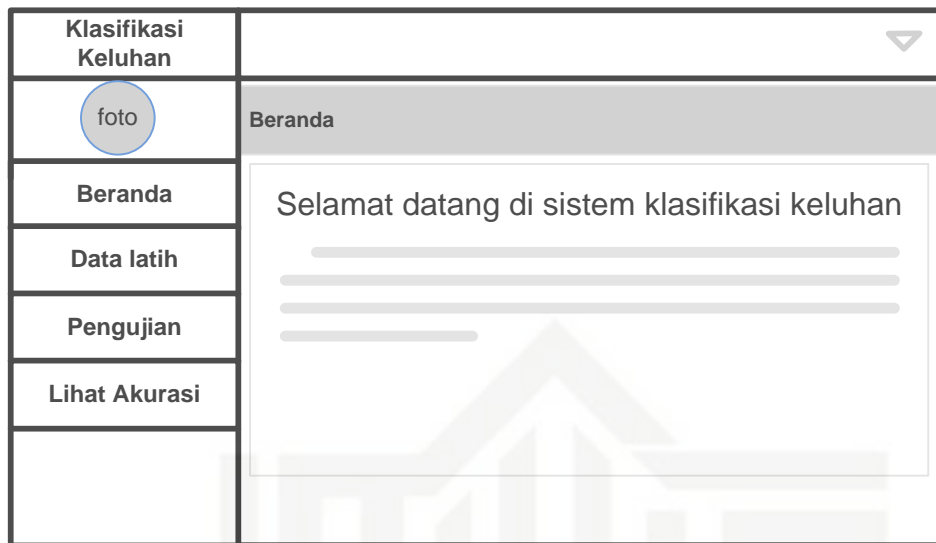
Interface sistem adalah Antarmuka (*Interface*) merupakan mekanisme komunikasi antara pengguna (*user*) dengan sistem. Antarmuka (*Interface*) dapat menerima informasi dari pengguna (*user*) dan memberikan informasi kepada pengguna (*user*) untuk membantu mengarahkan alur penelusuran masalah sampai ditemukan suatu solusi. Berikut tampilan *interface* sistem klasifikasi keluhan mahasiswa ini :

A. Tampilan Halaman Beranda

Halaman beranda adalah halaman utama yang ditampilkan aplikasi kepada pengguna. Halaman ini berisi tentang informasi sistem klasifikasi keluhan, alur kerja sistem, dan tujuan sistem. Berikut ini adalah Gambar 4.9 rancangan *interface* halaman beranda :

Hak Cipta Dilindungi Undang-Undang

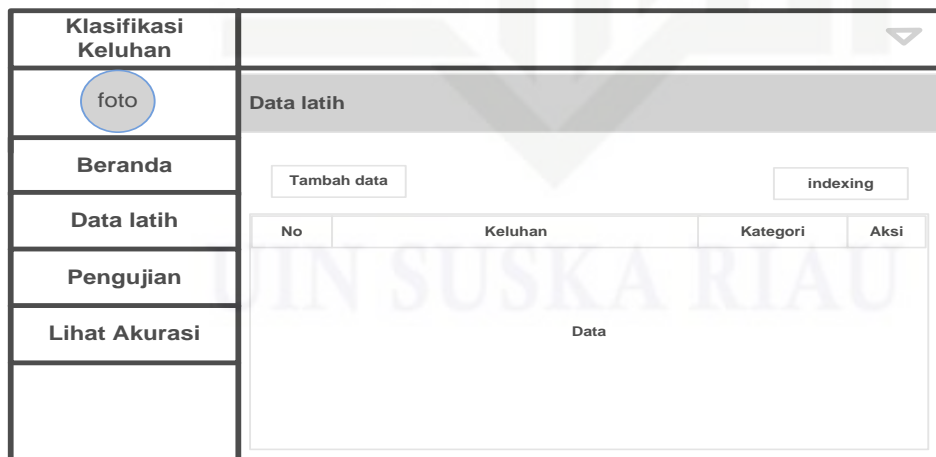
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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



Gambar 4.9 Rancangan *Interface* Halaman Beranda

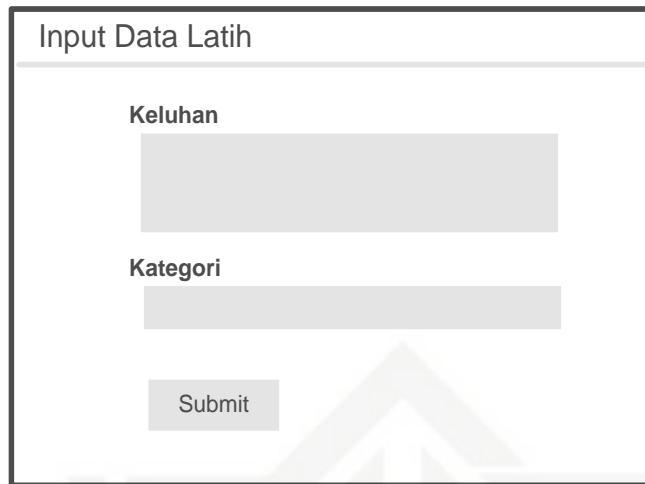
B. Tampilan Halaman Data Latih

Halaman data latih merupakan halaman yang memuat koleksi dokumen yang menjadi data latih pada sistem. Dalam halaman ini dapat melihat daftar dokumen keluhan juga terdapat tombol tambah data, indexing dan hapus dokumen. Tombol tambah data berfungsi untuk menampilkan form input tambah dokumen keluhan. Tombol *indexing* berfungsi untuk mengindex dokumen sedangkan *Link* hapus berfungsi untuk menghapus dokumen data latih. Berikut ini Gambar 4.10 adalah rancangan *interface* menu halaman latih:



Gambar 4.10 Rancangan *Interface* Halaman Latih

Berikut ini Gambar 4.11 adalah rancangan *interface* halaman tambah data latih:



Gambar 4.11 Rancangan *Interface* Halaman Tambah Data Latih

C. Tampilan Halaman Pengujian

Halaman pengujian merupakan halaman yang memuat dokumen data uji. Dalam halaman ini juga terdapat tombol tambah data dokumen, cari kategori, hapus dan *reset*. Tombol tambah data berfungsi menambah dokumen data uji. Tombol cari semua akan memberikan hasil pencarian kategori dari masing-masing keluhan berdasarkan metode yang digunakan. Sedangkan pencarian kategori secara detail dari tiap keluhan dapat dilihat pada kolom aksi pada tabel. Sedangkan tombol *reset* akan memberikan aksi mengembalikan nilai kategori keluhan pada keadaan awal. Berikut ini adalah Gambar 4.12 yaitu rancangan *interface* halaman pengujian :

Klasifikasi Keluhan	
foto	Pengujian
Beranda	Tambah data Cari semua Reset
Data latih	No Keluhan Kategori Aksi
Pengujian	
Lihat Akurasi	Data

Gambar 4.12 Rancangan *Interface* Halaman Pengujian

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.

D. Tampilan Halaman Lihat Akurasi

Halaman lihat akurasi akan memuat informasi hasil pengujian yang telah dilakukan. Terdapat informasi akurasi pengujian, Berikut halaman lihat akurasi dijelaskan Gambar 4.13

Klasifikasi Keluhan	
foto	Akurasi
Beranda	
Data latih	
Pengujian	
Lihat Akurasi	

Gambar 4.13 Rancangan Interface Halaman Lihat Akurasi