

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.

BAB IV

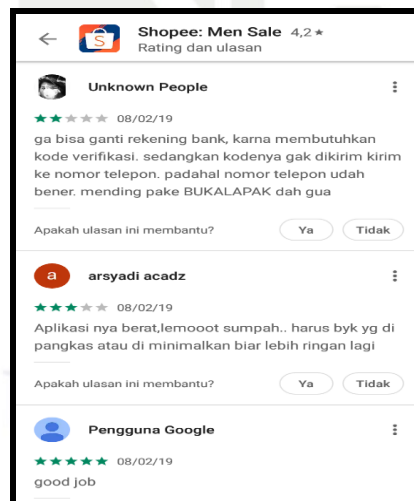
ANALISA DAN PERANCANGAN

4.1 Analisa

Analisa merupakan tahapan proses pengkajian yang sangat penting dalam melakukan suatu penelitian untuk membahas secara mendalam dari suatu pokok permasalahan pada penelitian ini. Pada bagian analisa berisikan pembahasan mengenai analisa permasalahan, analisa kebutuhan data, praproses data, pembobotan dan proses klasifikasi menggunakan metode *Naïve Bayes Classifier* untuk mengetahui tingkat akurasi.

4.1.1 Analisa Masalah

Permasalahan dalam penelitian ini bagaimana mengklasifikasikan pemberi komentar/ulasan pada aplikasi *Shopee* ke dalam 3 kelas yaitu positif, netral, dan negatif berdasarkan komentar/ulasan yang ada pada aplikasi *Google play* dengan proses klasifikasi menggunakan metode *Naïve Bayes Classifier*. Gambar 4.1 menampilkan komponen pada halaman. Shopee yang akan digunakan dalam penelitian.



Gambar 4.1 Komponen Pada Aplikasi Shopee



Hak Cipta Dilindungi Undang-Undang

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

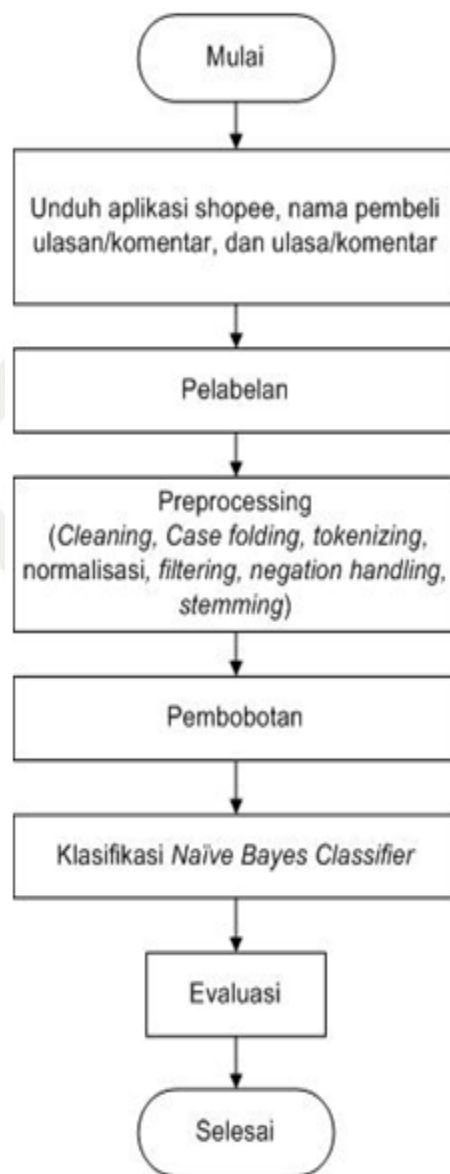
- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

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

Berikut adalah penjelasan dari komponen-komponen pada gambar 4.1:

1. Nama Aplikasi, digunakan untuk pengunduhan data komentator, pemberi komentar/ulasan pada Aplikasi Shopee.
2. Komentar, teks yang nantinya akan diklasifikasikan.

Adapun gambaran umum proses-proses sebagai berikut:



Gambar 4.2 Gambaran Umum Proses



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.

4.1.2 Analisa Kebutuhan Data

Analisa kebutuhan data digunakan untuk mengetahui data apa saja yang diperlukan dalam proses klasifikasi. Data yang dibutuhkan pada penelitian ini adalah nama pemberi komentar, dan teks komentar. Komentar Shopee yang digunakan sebagai data latih dan data uji dikumpulkan melalui Aplikasi Shopee Indonesia.

Dari nama-nama pemberi komentar pada aplikasi Shopee akan dipilih *dataset* sebanyak 1500 komentar menggunakan bahasa Indonesia setelah pelabelan manual. Pelabelan manual dilakukan oleh peneliti yang selanjutnya diperiksa oleh lulusan SI Jurusan Psikologi atas nama Rahma Tri Aulia, S.Psi. Klasifikasi terdiri dari 3 kelas dengan 500 komentar positif, 500 komentar netral, dan 500 komentar negatif. Perbandingan *dataset* yang digunakan adalah 80% data latih : 20% data uji. Tabel 4.1 berisikan rincian *dataset* yang digunakan:

Tabel 4. 1 Rincian Dataset

Data Komentar	Data Latih	1200 (80%)		
	Data Uji	300 (20%)		
		Latih	Uji	Jumlah
Kelas	Positif	400	100	500
	Negatif	400	100	500
	Netral	400	100	500
	Jumlah	1.200 komentar	300 Komentar	1.500 Komentar

Pengumpulan data dilakukan dengan cara *scraping* menggunakan ekstensi *googlechrome* yang bernama data miner dan terhubung ke internet. Untuk pengumpulan data dibutuhkan data dalam format CSV (*Comma Separated Value*) yang didapat dari aplikasi Shopee pada *Google play*.

Pada tabel 4.2 dibawah ini adalah contoh data komentar yang telah dikumpulkan dan diberi kelas.

Hak Cipta Dilindungi Undang-Undang

1. Diarangi 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.

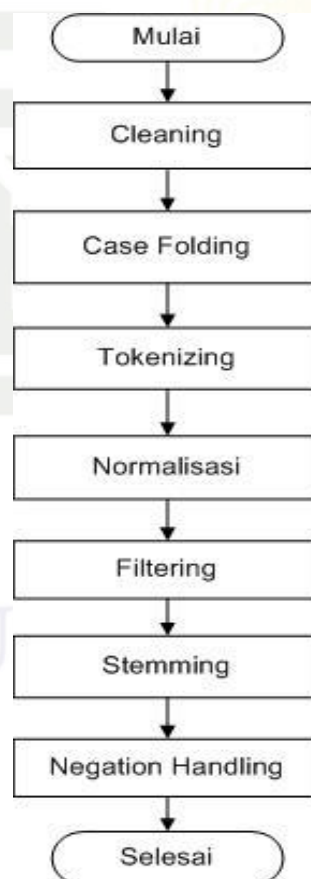
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 4. 2 Contoh Data Komentar

No	Komentator	Komentar	Kelas
1	muhammad hasan	bagus,,pokoknya shopee bgus	Positif
2	cahya seva	kecewa tingkat dewa!!	Negatif
3	bara hendrawanto	ka, 1 minggu berlalu aku beli produk dari @viasofid bahan bajunya ga jelek, makasih ya ka produknya bagus banget, recommended banget dejh ka	Netral

4.1.3 Text Preprocessing

Praproses (*Preprocessing*) atau tahapan praproses merupakan tahapan yang penting dalam melakukan analisa klasifikasi, bertujuan untuk membersihkan data-data dari unsur-unsur yang tidak dibutuhkan berguna untuk mempermudah dalam proses klasifikasi. Gambar 4.3 merupakan *flowchart* tahapan praproses data yang digunakan.



Gambar 4. 3 Flowchart Praproses Data

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 adalah penjelasan dari Gambar 4.3 flowchart praproses data yang digunakan:

1. *Cleaning*

Cleaning akan membersihkan atau menghilangkan karakter dan entitas pada komentar yang tidak diperlukan. Adapun kata atau karakter yang akan dihilangkan pada data komentar adalah karakter atau simbol (@#\$\$%^&*()_+”:{}<>.,?!~/[]), dan angka. Untuk lebih jelasnya dapat dilihat pada tabel 4.3 berikut:

Tabel 4. 3 Hasil *Cleaning* Data Komentar Shopee

No	Komentar	Hasil <i>Cleaning</i>
1	Bagus,, pokoknya shopee bagus	Bagus pokoknya shopee bagus
2	Kecewa tingkat dewa!!	Kecewa tingkat dewa
3	Ka, 1 minggu berlalu aku beli Produk dari @viasofid bahan bajunya ga jelek, Makasih ya ka produknya bagus banget, Recommended banget dejh ka?	Ka minggu berlalu aku beli Produk dari bahan bajunya ga jelek Makasih ya ka produknya bagus banget Rekomended banget dejh ka

2. *Case Folding*

Proses *case folding* adalah tahapan penyeragaman bentuk huruf dengan mengubah semua huruf yang ada menjadi huruf kecil. Tabel 4.4 adalah hasil data yang telah dilakukan proses *case folding*:

Tabel 4.4 Hasil *Case Folding* Data Komentar Shopee.

No	Komentar	Hasil <i>Case Folding</i>
1	Bagus pokoknya shopee bagus	bagus pokoknya shopee bagus
2	Kecewa tingkat dewa	kecewa tingkat dewa
3	Ka minggu berlalu aku beli Produk dari bahan bajunya ga jelek Makasih ya ka produknya bagus banget Rekomended banget dejh ka	ka minggu berlalu aku beli produk dari bahan bajunya ga jelek makasih ya ka produknya bagus banget rekomended banget dejh ka

3. *Tokenizing*

Proses *Tokenizing* yaitu tahapan dilakukannya pengenalan dan ekstraksi pada setiap kata. Bertujuan untuk proses memisahkan atau memecah teks atau kalimat menjadi sebuah kumpulan kata agar mudah dalam melakukan

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.

pembobotan tiap kata nantinya. Hasil *tokenizing* untuk contoh data komentar dapat dilihat pada tabel 4.5:

Tabel 4.5 Hasil *Tokenizing* Data Komentar Shopee

Data 1	Data 2	Data 3
bagus	kecewa	ka
pokoknya	tingkat	minggu
shopee	dewa	berlalu
bagus		aku
		beli
		produk
		dari
		bahan
		bajunya
		ga
		jelek
		makasih
		ya
		ka
		produknya
		bagus
		banget
		rekomended
		banget
		dejh
		ka

4. Normalisasi

Normalisasi merupakan proses yang bertujuan untuk mengubah kata yang tidak sesuai ejaan. Proses untuk menangani kata seperti ini dilakukan dengan melibatkan kamus yang dibuat terdiri dari kata tak baku dan kata bakunya.

Berikut tabel 4.6 adalah hasil data yang dilakukan proses normalisasi:

Tabel 4. 6 Hasil Normalisasi Data Komentar Shopee

Data 1	Data 2	Data 3
bagus	kecewa	kak
pokoknya	tingkat	minggu
shopee	dewa	berlalu
bagus		aku
		beli
		produk
		dari
		bahan

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.

	bajunya
	tidak
	jelek
	terima
	kasih
	ya
	kak
	produknya
	bagus
	banget
	rekomended
	banget
	deh
	kak

5. Filtering

Proses *filtering (stopward)* merupakan tahapan untuk menghilangkan kata-kata yang kurang penting yang dikhawatirkan akan mengganggu proses klasifikasi. Untuk kata yang akan di *stopword* menggunakan kamus dari KBBI V online dipilih berdasarkan jenis katanya yaitu kata keterangan (*adverbia*), kata ganti (pronomina), kata seru (interjeksi), kata depan (preposisi), dan kata hubung (konjungsi). Tabel 4.7 adalah hasil data yang telah dilakukan proses *filtering*:

Tabel 4. 7 Hasil Filtering Data Komentar Shopee

Data 1	Data 2	Data 3
bagus	kecewa	kak
shopee	tingkat	minggu
bagus	dewa	berlalu
		beli
		produk
		bahan
		bajunya
		tidak
		jelek
		terima
		kasih
		kak
		produknya
		bagus
		kak

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.

6. Stemming

Proses *stemming* merupakan proses untuk mengubah kata-kata, akhiran yang tidak mengubah kata dasar *Inflection Suffixes (IS)*, akhiran yang secara langsung ditambahkan pada kata dasar *Derivation Suffixes (DS)*, serta awalan yang dapat langsung diberikan pada kata dasar murni *Derivation Prefixes (DP)*, dengan menggunakan algoritma *Enhanced Confix Stripping (ECS)*. Proses stemming dengan Algoritma *Enhanced Confix Stripping (ECS)* ini telah digambarkan sebelumnya pada bagian kedua. Langkah-langkah pada *stemming* menggunakan ECS ini yaitu dengan mengembalikan sebuah kata menjadi bentuk dasar dengan menghilangkan imbuhan, jika kata terdapat didalam kamus maka dikembalikan ke kata dasar. Namun, jika kata tidak ada didalam kamus, maka kata tidak dikembalikan. Berikut tabel 4.9 adalah hasil data yang dilakukan proses *stemming*.

Tabel 4. 8 Hasil Stemming Data Komentar Shopee

Data 1	Data 2	Data 3
bagus	kecewa	kak
shopee	tingkat	minggu
bagus	dewa	lalu
		beli
		produk
		bahan
		baju
		bagus
		terima
		kasih
		kak
		produk
		bagus
		kak

7. Negation Handling

Negation Handling atau penanganan negasi merupakan mencari kata yang bersifat negasi, karena kata negasi dapat mengubah nilai positif atau negatif dari data. dengan cara mendeteksi kata negasi dan membalikkan kata setelah kata negasi tersebut dengan cara menggunakan kamus. Contohnya seperti kata

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.

“tidak baik” menjadi “buruk”. Berikut adalah hasil data yang telah dilakukan proses negation handling.

Tabel 4. 9 Hasil Negation Handling Data Komentar Shopee

Data 1	Data 2	Data 3
bagus	kecewa	kak
shopee	tingkat	minggu
bagus	dewa	berlalu
		beli
		produk
		bahan
		bajunya
		bagus
		terima
		kasih
		kak
		produknya
		bagus
		kak

4.1.1 Pembobotan

Pada penelitian ini digunakan pembobotan dengan menggunakan *tf* (*term frequency*) Pembobotan merupakan proses merubah kata menjadi bentuk angka atau *vector*. TF merupakan jumlah kemunculan kata dalam dokumen. Kata direpresentasi ke dalam bentuk *vector*, dimana tiap kata dihitung sebagai satu fitur.

Adapun perhitungan bobot yang digunakan merupakan *Term Frequency* (TF). Pada tabel 4.10 adalah hasil pembobotan selengkapnya berdasarkan data komentar di tabel sebelumnya.

Tabel 4. 10 Hasil Pembobotan Kata

No	Kosa Kata	<i>tf(pos)</i>	<i>tf(neg)</i>	<i>tf(net)</i>
1	bagus	2	0	2
2	shopee	1	0	0
3	kecewa	0	1	0
4	tingkat	0	1	0
5	dewa	0	1	0
6	kak	0	0	3
7	minggu	0	0	1
8	lalu	0	0	1
9	beli	0	0	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.

10	produk	0	0	2
11	bahan	0	0	1
12	baju	0	0	1
13	terima	0	0	1
14	kasih	0	0	1

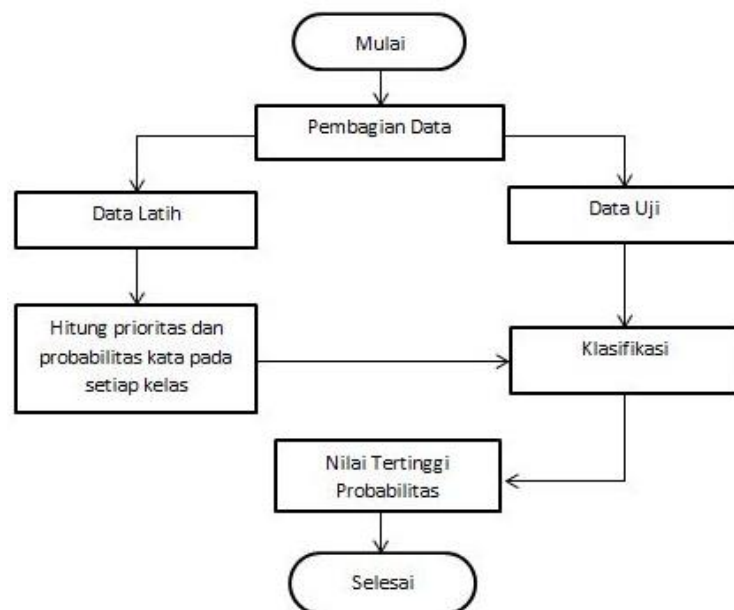
Keterangan:

Kosa kata : Kata tunggal yang telah terpisah dari dokumen query

$tf(n)$: *Term Frequency* (frekuensi kata) pada dokumen ke- n

4.1.5. Proses Klasifikasi Metode *Naïve Bayes Classifier*

Berdasarkan analisa kebutuhan data dan tahap praproses yang telah dijabarkan sebelumnya, maka pada tahapan ini akan dijelaskan teknik-teknik yang akan digunakan dalam pengklasifikasian data. *Flowchart* proses *Naïve Bayes* dalam klasifikasi penelitian ini akan diperlihatkan pada gambar 4.4.



Gambar 4.4 Flowchart Proses Klasifikasi

Pada proses klasifikasi *Naïve Bayes* ini, seluruh dataset dibagi menjadi beberapa tahap data *training* (latih) dan data *testing* (uji). Berikut adalah penjelasan langkah-langkahnya beserta contoh perhitungan:



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.

1. Tahap data *training* (latih)

Pada tahap training ini data yang telah didapatkan nilai bobotnya seperti contoh pada tabel 4.10 dijadikan sebagai data latih untuk menjadi acuan atau petunjuk dalam membentuk model klasifikasi. Pada tahap ini akan dicari nilai *prioritas* probabilitas kategori dan *conditional probabilities* masing-masing kata dari data latih. Menghitung nilai *prioritas* dan *conditional probabilities* persamaan 2.1 dan persamaan 2.2) setiap *term* untuk setiap kelas dari data latih. Perhitungannya adalah sebagai berikut:

- Pertama hitung probabilitas setiap kategori (*prioritas*) menggunakan persamaan 2.1, dalam penelitian ini yang menjadi kategori ada tiga kategori yaitu kelas positif, kelas netral, dan kelas negatif.

$$P(\text{positif/negatif/netral}) = \frac{D(\text{positif/negatif/netral})}{C}$$

$$p(\text{positif}) = \frac{f_d(\text{positif})}{|C|} = \frac{1}{3}$$

$$p(\text{negatif}) = \frac{f_d(\text{negatif})}{|C|} = \frac{1}{3}$$

$$p(\text{netral}) = \frac{f_d(\text{netral})}{|C|} = \frac{1}{3}$$

- Setelah melakukan perhitungan probabilitas setiap kategori kemudian selanjutnya menghitung probabilitas setiap *term* dari semua dokumen dengan menggunakan persamaan 2.2. jumlah seluruh *term* yang akan digunakan dalam perhitungan ini adalah sebanyak 20 *term*, 3 *term* dari kelas positif, 3 *term* dari kelas negatif, dan 14 *term* dari kelas netral. Banyaknya *term* tergantung pada hasil praproses data. Berdasarkan contoh yang ada pada tabel 4.9, perhitungan probabilitas dari setiap *term* 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.

$$P(w|\text{positif/negatif/netral}) = \frac{(n_k, \text{positif/negatif/netral}) + 1}{(n, \text{positif/negatif/netral}) + |T|}$$

Diketahui $|T| = 20$

count positif = 3, *count* negatif = 3, *count* netral = 14

i. Probabilitas kata “bagus”

$$p(\text{"bagus"}|\text{"pos"}) = \frac{(f(\text{"bagus"}|\text{"pos"})+1)}{f(\text{"pos"})+|T|} = \frac{2+1}{3+20} = 0.130434783$$

$$p(\text{"bagus"}|\text{"neg"}) = \frac{(f(\text{"bagus"}|\text{"neg"})+1)}{f(\text{"neg"})+|T|} = \frac{0+1}{3+20} = 0.043478260$$

$$p(\text{"bagus"}|\text{"net"}) = \frac{(f(\text{"bagus"}|\text{"net"})+1)}{f(\text{"net"})+|T|} = \frac{2+1}{14+20} = 0.0882352941$$

ii. Probabilitas kata “shopee”

$$p(\text{"shopee"}|\text{"pos"}) = \frac{(f(\text{"shopee"}|\text{"pos"})+1)}{f(\text{"pos"})+|T|} = \frac{1+1}{3+20} = 0.0869565217$$

$$p(\text{"shopee"}|\text{"neg"}) = \frac{(f(\text{"shopee"}|\text{"neg"})+1)}{f(\text{"neg"})+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p(\text{"shopee"}|\text{"net"}) = \frac{(f(\text{"shopee"}|\text{"net"})+1)}{f(\text{"net"})+|T|} = \frac{0+1}{14+20} = 0.0294117647$$

iii. Probabilitas kata “kecewa”

$$p(\text{"kecewa"}|\text{"pos"}) = \frac{(f(\text{"kecewa"}|\text{"pos"})+1)}{f(\text{"pos"})+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p(\text{"kecewa"}|\text{"neg"}) = \frac{(f(\text{"kecewa"}|\text{"neg"})+1)}{f(\text{"neg"})+|T|} = \frac{1+1}{3+20} = 0.0869565217$$

$$p(\text{"kecewa"}|\text{"net"}) = \frac{(f(\text{"kecewa"}|\text{"net"})+1)}{f(\text{"net"})+|T|} = \frac{0+1}{14+20} = 0.0294117647$$

iv. Probabilitas kata “tingkat”

$$p(\text{"tingkat"}|\text{"pos"}) = \frac{(f(\text{"tingkat"}|\text{"pos"})+1)}{f(\text{"pos"})+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p(\text{"tingkat"}|\text{"neg"}) = \frac{(f(\text{"tingkat"}|\text{"neg"})+1)}{f(\text{"neg"})+|T|} = \frac{1+1}{3+20} = 0.0869565217$$

$$p(\text{"tingkat"}|\text{"net"}) = \frac{(f(\text{"tingkat"}|\text{"net"})+1)}{f(\text{"net"})+|T|} = \frac{0+1}{14+20} = 0.0294117647$$

v. Probabilitas kata “dewa”

$$p(\text{"dewa"}|\text{"pos"}) = \frac{(f(\text{"dewa"}|\text{"pos"})+1)}{f(\text{"pos"})+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p(\text{"dewa"}|\text{"neg"}) = \frac{(f(\text{"dewa"}|\text{"neg"})+1)}{f(\text{"neg"})+|T|} = \frac{1+1}{3+20} = 0.0869565217$$

$$p(\text{"dewa"}|\text{"net"}) = \frac{(f(\text{"dewa"}|\text{"net"})+1)}{f(\text{"net"})+|T|} = \frac{0+1}{14+20} = 0.0294117647$$



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.

vi. Probabilitas kata “kak”

$$p("kak"|"pos") = \frac{(f("kak"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("kak"|"neg") = \frac{(f("kak"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("kak"|"net") = \frac{(f("kak"|"net")+1)}{f("net")+|T|} = \frac{3+1}{14+20} = 0.117647059$$

vii. Probabilitas kata “minggu”

$$p("minggu"|"pos") = \frac{(f("minggu"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("minggu"|"neg") = \frac{(f("minggu"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("minggu"|"net") = \frac{(f("minggu"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

viii. Probabilitas kata “lalu”

$$p("lalu"|"pos") = \frac{(f("lalu"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("lalu"|"neg") = \frac{(f("lalu"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("lalu"|"net") = \frac{(f("lalu"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

ix. Probabilitas kata “beli”

$$p("beli"|"pos") = \frac{(f("beli"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("beli"|"neg") = \frac{(f("beli"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("beli"|"net") = \frac{(f("beli"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

x. Probabilitas kata “produk”

$$p("produk"|"pos") = \frac{(f("produk"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("produk"|"neg") = \frac{(f("produk"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("produk"|"net") = \frac{(f("produk"|"net")+1)}{f("net")+|T|} = \frac{2+1}{14+20} = 0.0882352941$$

xi. Probabilitas kata “bahan”

$$p("bahan"|"pos") = \frac{(f("bahan"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("bahan"|"neg") = \frac{(f("bahan"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

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.

$$p("bahan"|"net") = \frac{(f("bahan"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

xii. Probabilitas kata “baju”

$$p("baju"|"pos") = \frac{(f("baju"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("baju"|"neg") = \frac{(f("baju"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("baju"|"net") = \frac{(f("baju"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

xiii. Probabilitas kata “terima”

$$p("terima"|"pos") = \frac{(f("produk"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("terima"|"neg") = \frac{(f("produk"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("terima"|"net") = \frac{(f("produk"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

xiv. Probabilitas kata “kasih”

$$p("kasih"|"pos") = \frac{(f("produk"|"pos")+1)}{f("pos")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("kasih"|"neg") = \frac{(f("produk"|"neg")+1)}{f("neg")+|T|} = \frac{0+1}{3+20} = 0.0434782609$$

$$p("kasih"|"net") = \frac{(f("produk"|"net")+1)}{f("net")+|T|} = \frac{1+1}{14+20} = 0.0588235294$$

2. Tahapan *testing* (uji)

Pada tahapan *testing* (uji) ini, akan dilakukan pengujian dengan memasukkan data uji ke dalam model yang telah dibentuk pada tahap *training* tadi. Pada table 4.11 adalah contoh data yang akan dijadikan sebagai data uji.

Tabel 4. 11 Contoh Data Uji

Komentator	Komentator	Kelas
@viasofid	ka produk bahan baju dari @viasofid aku uda beli loh ternyata beneran hasilnya bagus dengan gambar info hubungi watshapp : 081318804085	Netral

Berdasarkan data uji dari tabel 4.11 di atas, maka akan dihitung nilai probabilitasnya berdasarkan nilai probabilitas setiap *term*. Perhitungan nilai probabilitas dengan metode *Naïve Bayes* pada dokumen *testing* ini dilakukan

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.

dengan mengalikan nilai probabilitas semua kategori dengan probabilitas tiap *term* yang diambil dari semua data latih.

Sebelum masuk pada tahap perhitungan selanjutnya, data terlebih dahulu melewati *preprocessing*. Hasil data testing setelah di *preprocessing* yaitu seperti pada tabel 4.12.

Tabel 4. 12 Data Uji Setelah Praproses

Komentator	Komentator	Kelas
@viasofid	kak produk bahan baju beli nyata hasil bagus gambar info hubungi	Netral

Selanjutnya kita akan melakukan *conditional probabilities* pada data uji dengan menggunakan hasil probabilitas term yang di dapat pada data latih, dimana pada data latih dan data uji melakukan tahapan *preprocessing* seperti tabel 4.13 berikut:

Tabel 4. 13 Contoh Tabel Data Latih Data Uji

Data Latih (n)	Komentar	Kelas
1	bagus semua shopee bagus	Positif
2	kecewa tingkat dewa	Negatif
3	kak minggu lalu beli produk bahan baju bagus terima kasih kak produk bagus kak	Netral
Data Uji		
4	kak produk bahan baju beli nyata hasil bagus gambar info hubungi	?

Conditional probalities dilakukan menggunakan rumus 2.2

$$P(w|\text{positif/negatif/netral}) = \frac{(n_k, \text{positif/negatif/netral}) + 1}{(n, \text{positif/negatif/netral}) + |T|}$$

Diketahui $|T| = 20$

count positif = 3, *count* negatif= 3, *count* netral= 14

Kata data uji “kak produk bahan baju beli nyata hasil bagus gambar info hubungi”

$$P(\text{kak} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{produk} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{bahan} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{baju} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

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.

$$P(\text{beli} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{nyata} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{hasil} | \text{pos}) = (0+!)/(3+20) = 0.0434782609$$

$$P(\text{bagus} | \text{pos}) = (2+1)/(3+20) = 0.130434783$$

$$P(\text{gambar} \mid \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{info} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{hubungi} | \text{pos}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{kak} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{produk} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{bahan} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{baju} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{beli} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{nyata} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{hasil} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{bagus} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{gambar} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{info} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{hubungi} \mid \text{neg}) = (0+1)/(3+20) = 0.0434782609$$

$$P(\text{kak} \mid \text{net}) = (3+1)/(14+20) = 0.117647059$$

$$P(\text{produk} | \text{net}) = (2+1)/(14+20) = 0.0882352941$$

$$P(\text{bahan} \mid \text{net}) = (1+1)/(14+20) = 0.0588235294$$

$$P(\text{baju} \mid \text{net}) = (1+1)/(14+20) = 0.0588235294$$

$$P(\text{beli} \mid \text{net}) = (1+1)/(14+20) = 0.0588235294$$

$$P(\text{nyata} \mid \text{net}) = (0+1)/(14+20) = 0.0294117647$$

$$P(\text{hasil} \mid \text{net}) = (0+1)/(14+20) = 0.0294117647$$

$$P(\text{bagus} \mid \text{net}) = (2+1)/(14+20) = 0.130434783$$

$$P(\text{gambar} \mid \text{net}) = (0+1)/(14+20) = 0.0294117647$$

$$P(\text{info} \mid \text{net}) = (0+1)/(14+20) = 0.0294117647$$

$$P(\text{hubungi} \mid \text{net}) = (0+1)/(14+20) = 0.0294117647$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

Be 2.2

$$v(v) = 1$$

pos

$$v(v) = 1$$

neg

$$v(v) = 1$$

net

$$v(v) = 1$$

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.

$p(uji|netral)$

$$= 0.333333333 \times 0.117647059 \times 0.0882352941 \times 0.0588235294 \times 0.0588235294 \times 0.0588235294 \times 0.0294117647 \times 0.0294117647 \times 0.130434783 \times 0.0294117647 \times 0.0294117647 \times 0.0294117647 \\ = 2.02187465E-15$$

Nilai probabilitas tertinggi yaitu sebesar 2.02187465E-15 pada $p(uji|net)$, sehingga komentar tersebut diklasifikasikan kedalam kelas “Netral”.

4.2 Perancangan

Perancangan klasifikasi positif dan negatif berdasarkan komentar Shopee dirancang berdasarkan tahapan analisa yang telah dilakukan sebelumnya. Adapun bagian perancangan dilakukan sebagai acuan dalam pembuatan aplikasi di penelitian ini, yang terdiri dari perancangan basis data (*database*), perancangan struktur menu dan perancangan antar muka (*interface*).

4.2.1 Perancangan Basis Data (*Database*)

Perancangan basis data yang akan digunakan pada shopee terdiri dari 4 tabel, yaitu tabel kamus kata dasar (*basicword*), tabel kata henti (*stopword*), tabel kamus kata normalisasi, dan data Shopee (komentar). Berikut ini tabel pada *database* Shopee:

Tabel 4. 14 Nama Tabel *Database* Shopee

No	Nama Tabel
1	<i>Basicword</i>
2	<i>Stopword</i>
3	<i>Normalisasi</i>
4	<i>Komentar</i>
5	<i>Kata negasi</i>
6	<i>Negasi</i>
7	<i>Dataset</i>

Berikut adalah struktur pada tabel *basicword* atau kamus kata dasar yang terdiri dari 2 *field*, seperti yang terlihat pada tabel 4.15:

No	Nama Field	Type dan Length	Primary Key
1	Name	varchar (70)	YES
2	Type	varchar (25)	-

Tabel 4. 16 Struktur Tabel *Stopword*

No	Nama Field	Type dan Length	Primary Key
1	Name	varchar (70)	YES
2	Type	varchar (25)	-

Tabel 4. 17 Normalisasi

No	Nama Field	Type dan Length	Primary Key
1	Kata_tidakbaku	varchar (50)	YES
2	Kata_baku	varchar (50)	-

Tabel 4. 18 Struktur Tabel Komentar

No	Nama Field	Type dan Length	Primary Key
1	Id (<i>Utama</i>)	int (5)	YES
2	Username	varchar (50)	-
3	Komentar	varchar (50)	-
4	Preprocess	varchar (5000)	-
5	Kelas	varchar (5000)	-
6	Ket	varchar (50)	-

Tabel 4. 19 katanegasi

No	Nama Field	Type dan Length	Primary Key
1	Kata	varchar (50)	YES

Tabel 4. 20 negasi

No	Nama Field	Type dan Length	Primary Key
1	KataPositif (<i>Utama</i>)	varchar (50)	YES
2	KataNegatif	varchar (50)	-

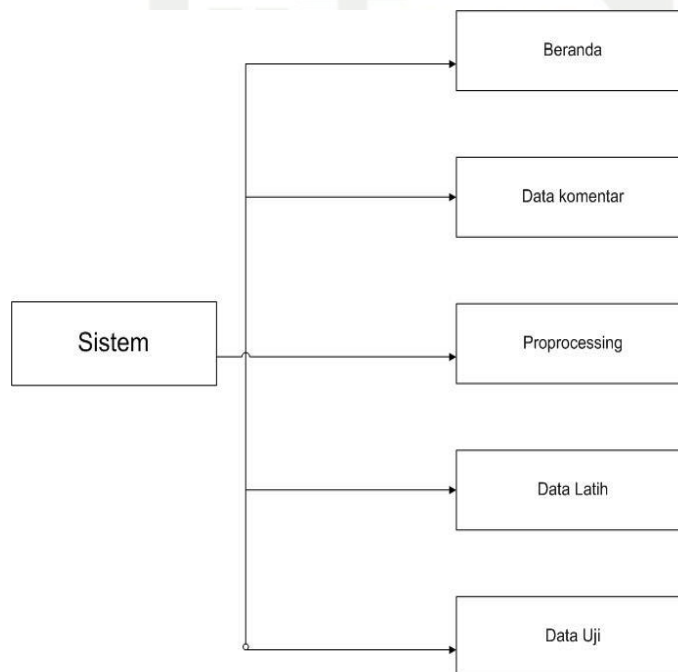
Berikut adalah struktur pada tabel dataset atau tabel data shopee terdiri dari 2 *field*, seperti yang terlihat pada tabel 4.21:

Tabel 4. 21 dataset

No	Nama Field	Type dan Length	Primary Key
1	data_latih	int(11)	YES
2	data_uji	int(11)	-

4.2.2 Perancangan Struktur Menu

Perancangan struktur menu menggambarkan hubungan halaman dengan halaman yang lainnya. Menu merupakan salah satu bagian penting dalam merancang sebuah aplikasi antar muka karena melalui menu dapat dilihat bagaimana struktur aplikasi antar muka ini terbentuk. Rancangan struktur menu dapat dilihat pada gambar 4.5 berikut:



Gambar 4.5 Rancangan Struktur Menu

4.2.3 Perancangan Antarmuka (*Interface*)

Perancangan Antarmuka aplikasi merupakan sarana pengembangan yang digunakan untuk membuat komunikasi yang baik antara aplikasi dengan penggunanya. Antarmuka dapat menerima informasi dari pengguna (*user*) dan memberikan informasi kepada pengguna (*user*) untuk membantu mengarahkan

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.

alur penelusuran masalah sampai ditemukan suatu solusi. Berikut merupakan tampilan antarmuka yang akan dirancang dalam penelitian ini.

1. Rancangan Antarmuka Halaman Beranda

Halaman beranda merupakan halaman utama yang akan diakses dan ditampilkan aplikasi kepada pengguna. Halaman ini berisi tentang informasi judul penelitian, nama dan nim. Berikut ini adalah Gambar 4.6 rancangan antar muka halaman beranda:



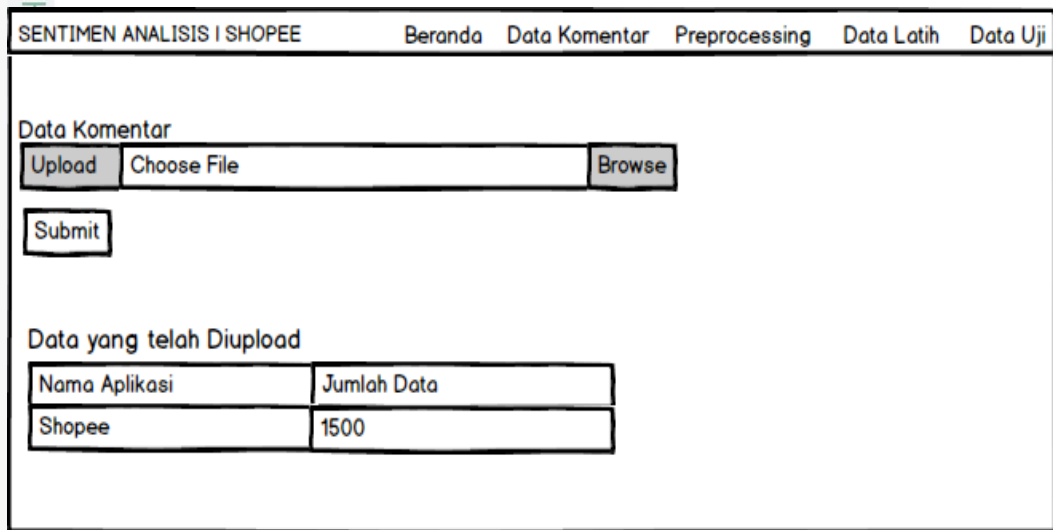
Gambar 4.6 Rancangan Antarmuka Halaman Beranda

2. Tampilan Halaman Unduh Data Shopee

Halaman unduh data merupakan halaman untuk mengunduh data-data yang akan digunakan untuk proses klasifikasi. Pada halaman ini juga terdapat tabel nama aplikasi dan jumlah data yang sudah di unduh. Berikut ini adalah Gambar 4.7 rancangan antar muka 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.

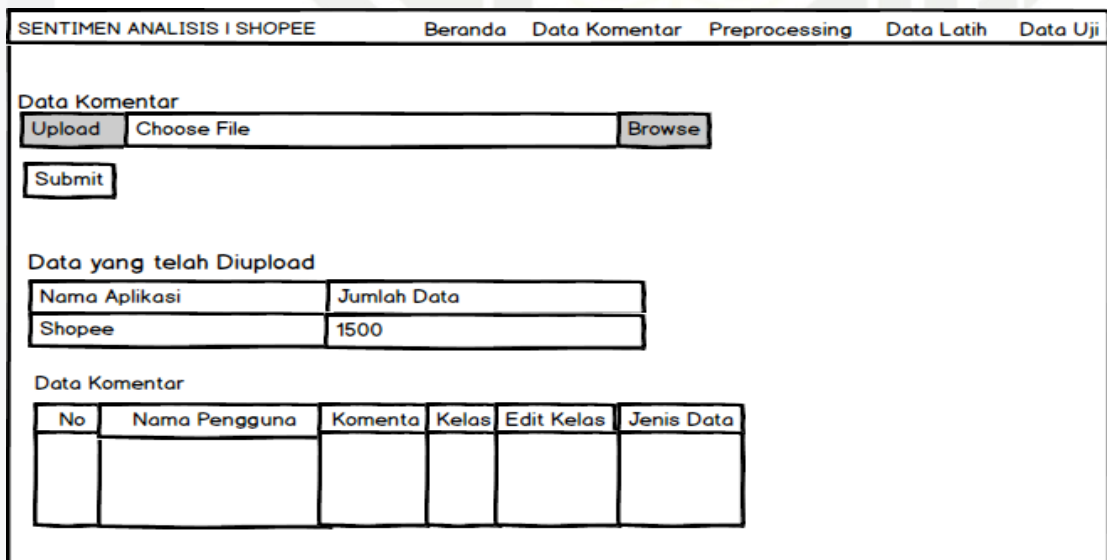


Nama Aplikasi	Jumlah Data
Shopee	1500

Gambar 4.7 Rancangan Antarmuka Halaman Unduh Data Shopee

3. Tampilan Halaman Data Shopee

Halaman data Shopee merupakan halaman yang menampilkan data yang sudah di unduh dan diberi label. Terdapat tombol Praproses untuk memulai *pre-processing* data. Berikut ini adalah Gambar 4.8 rancangan antar muka halaman beranda:



No	Nama Pengguna	Komenta	Kelas	Edit Kelas	Jenis Data

Gambar 4. 8 Rancangan Antarmuka Halaman Unduh Data Shopee

Pada halaman Pre-proses menampilkan tabel hasil praproses data yang berisikan nama komentator, komentarnya, dan hasil preprocess.

SENTIMEN ANALISIS | SHOPEE Beranda Data Komentar Preprocessing Data Latih Data Uji

Pre-Proses Search

No		Username	Komentar	Preprocess

Gambar 4.9 Rancangan Antarmuka Halaman Preproses

4. Tampilan Halaman Data Latih

Halaman data latih merupakan halaman yang menampilkan data yang sudah di praproses tadi untuk selanjutnya dilakukan proses data latih menggunakan metode *Naïve Bayes* yang tombol prosesnya terletak diatas tabel. Pilihan proses data latih terdiri dari 3 yaitu untuk proses data latih 70%, 80% dan 90%.

SENTIMEN ANALISIS | SHOPEE Beranda Data Komentar Preprocessing Data Latih Data Uji

Learning Rate

Persentase Data Latih Jumlah Data Latih

Persentase Data Uji Jumlah Data Uji

Simpan

Data	Persen (%)	Jumlah
Data Latih		
Data Uji		

Data	Positif	Negatif	Netral
Data Latih			
Data Uji			

Gambar 4.10 Rancangan Antarmuka Halaman Data Latih



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.

5. Tampilan Halaman Data Uji

Halaman data uji merupakan halaman yang menampilkan data yang akan diuji untuk selanjutnya dilakukan proses data uji yang tombol prosesnya terletak diatas tabel.

SENTIMEN ANALISIS I SHOPEE

Beranda Data Komentar Preprocessing Data Latih Data Uji

Proses Data Uji

Search

No	Nama	komentar	kelas

Gambar 4.11 Rancangan Antarmuka Halaman Data Uji