

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data ulasan aplikasi *Threads, an Instagram App* yang terdapat pada situs *Google Play Store*. Data dikumpulkan dengan melakukan teknik *web scraping* menggunakan perangkat lunak *Google Colab* dengan bantuan *library google-play-scraper* yang tersedia. *Scraping* dilakukan dengan mengambil *id* dari aplikasi *Threads*, yaitu *com.instagram.barcelona*, kemudian data yang diambil merupakan ulasan yang ditulis dengan Bahasa Indonesia sejak aplikasi diluncurkan (5 Juli 2023) hingga 9 Februari 2024. Data yang berhasil dikumpulkan sebanyak 5572 yang kemudian disaring kembali secara manual sehingga menghasilkan 3322 data. Data yang tidak terpilih adalah ulasan yang ditulis dengan Bahasa Inggris secara keseluruhan dan ulasan yang tidak memiliki keterhubungan dengan aplikasi, sehingga tidak memberikan sentimen positif atau negatif. Data yang telah diterima kemudian disimpan dalam format *.csv*. Hasil *scraping* dapat dilihat pada Gambar 4.1.

source	review_id	user_name	review_title	review_description	rating	thumbs_up	review_date	developer_response	developer_response_date	appVersion	language_code	country_code
Google Play	3f598755-4999-4915-90f3-d8f3a1372dcd	Dukur Baraya	None	Ok	5	0	2024-02-01 19:09:26	None	None	300.0.0.29.109	id	id
Google Play	868ad514-9100-47d8-ae96-4091930076d	Muhammad Rizki Bismantara bintang kejora	None	Mantap	5	0	2024-01-22 15:13:48	None	None	None	id	id
Google Play	199f150c-a9aa-43a4-82dc-894658463c	Bj.jhon	None	Kasih bintang★★★★★ dulu nanti baru liat hasilnya	5	0	2024-01-21 11:59:32	None	None	None	id	id
Google Play	8a667201-539c-4656-9944-17c909b78ca7	Andhi Ti Ratumanan	None	Ok	5	0	2024-01-21 02:11:31	None	None	None	id	id
Google Play	35bee8c1-c89c-481f-9718-a1476d9d915	Love Hunter	None	apk sangat bagus, saya banyak melihat cewe labrut	5	1	2024-01-19 15:09:16	None	None	None	id	id
...	...	...	...	...	...	...	...	...	...	...	...	...
Google Play	9c86026f-7b33-4687-a61f-074aaa13b5a	Rehnn	None	selesai tuh Twitter -mark Zuckerberg	5	10	2023-07-06 01:24:12	None	None	289.0.0.77.109	id	id
Google Play	40d2b32-2605-4624-980c-02d4d586a76d	Drajat Achmad Inransyah A	None	Ini aplikasi, pas saya mau upload photo kok il..	3	15	2023-07-06 01:18:20	None	None	289.0.0.77.109	id	id
Google Play	93958a2f-7365-417a-a5cf-33f7c95d2224	Wahyyv	None	Apakah saya orang Indo pertama yang pakai Thre...	5	0	2023-07-06 01:18:12	None	None	None	id	id
Google Play	79695c7b-b338-4848-a80c-3578b33800b	Soleha Soleha	None	We hope that in the next update, a translate f...	5	1	2023-07-06 01:18:01	None	None	289.0.0.77.109	id	id
Google Play	1fa36806-16c0-4e6e-acdf-536fa00d5f	Willy	None	Kurang Tau	5	0	2023-07-06 01:08:30	None	None	289.0.0.77.109	id	id

5572 rows × 13 columns

**Gambar 4.1 Hasil Scraping Data**

#### 4.2 Pelabelan Data

Label data dalam penelitian ini dibagi menjadi dua, yaitu data yang bersentimen positif dan negatif. Sebuah ulasan diberi label positif jika memuat pujian dan dukungan terhadap aplikasi, sementara dianggap negatif jika berisi kritik dan keluhan terhadap aplikasi. Pelabelan data ini dilakukan secara manual oleh tiga *annotator* yang ketiganya memiliki latar belakang ilmu komputer, dimana masing-masing *annotator* akan membaca tiap-tiap data ulasan yang ada,

kemudian memberikan label positif atau negatif sesuai dengan penilaiannya. Contoh proses pelabelan dapat dilihat pada tabel 4.1.

**Tabel 4.1 Contoh Pelabelan Data**

<i>Review</i>	<i>Annotator</i> 1	<i>Annotator</i> 2	<i>Annotator</i> 3	<i>Sentimen</i>
suka banget, bisa jadi tempat curhat dan catatan kecil hidupku 😊😊😊	Positif	Positif	Positif	<b>Positif</b>
Sebelumnya aplikasinya bagus simpn hk neko" dan cenderung sedikit bot karena login lewat Instagram. Tapi tolong di tambahkan gitu DM karena jadi gk perlu buka ig. Dan ada bug ketika di alihkan ke ig gk bisa kembali harus di tutup atau gk kembali ke home screen	Negatif	Positif	Negatif	<b>Negatif</b>
Keren aplikasinya. Coba bisa dipercepat lagi waktu utk posting kontennya pasti lebih seru 😊😊	Positif	Positif	Positif	<b>Positif</b>
Pertama kali log in langsung mengalami bug , tulisan tulisan tidak jelas , tampilan seperti layar handphone rusak , tolong cepat diperbaiki	Negatif	Negatif	Negatif	<b>Negatif</b>

Pada tabel 4.1 dapat dilihat bahwa terdapat perbedaan pendapat antara *annotator* ketika melabelkan ulasan. Hasil akhir label ditetapkan berdasarkan suara terbanyak, seperti pada ulasan kedua dan keempat, karena dua *annotator* menilai ulasan tersebut positif dan satu *annotator* menilai negatif, maka label

akhir ulasan tersebut adalah positif. Hasil akhir pelabelan data dapat dilihat pada tabel 4.2.

**Tabel 4.2 Hasil Pelabelan Data**

Sentimen	Positif	Negatif
Jumlah	1548	1774
Total	3322	

### 4.3 Data Pre-processing

Untuk dapat dilakukan klasifikasi, data masih belum siap untuk digunakan karena masih banyak mengandung noise atau kesalahan yang nantinya dapat mempengaruhi kinerja algoritma. Untuk itu, diperlukan tahapan *preprocessing* untuk membersihkan data dan memperbaiki kesalahan yang ada pada data. Dalam penelitian ini, *pre-processing* dilakukan dengan lima tahapan, yaitu *data cleansing*, *normalization*, *stopword removal*, *stemming*, dan *tokenizing* yang dilakukan dengan bahasa pemrograman *Python* dan bantuan *library* yang tersedia. Tabel 4.3 dibawah ini adalah contoh data yang belum melalui tahapan *preprocessing*.

**Tabel 4.3 Contoh Data sebelum Preprocessing**

suka banget, bisa jadi tempat curhat dan catatan kecil hidupku 😊😊😊
Sebemer nya aplikasinya bagus simpnhk neko" dan cenderung sedikit bot karena login lewat Instagram. Tapi tolong di tambahkan gitu DM karena jadi gk perlu buka ig. Dan ada bug ketika di alihkan ke ig gk bisa kembali harus di tutup atau gk kembali ke home screen
Keren aplikasinya. Coba bisa dipercepat lagi waktu utk posting kontennya pasti lebih seruuuu 👍👍
Pertama kali log in langsung mengalami bug , tulisan tulisan tidak jelas , tampilan seperti layar handphone rusak , tolong cepat diperbaiki

#### 4.3.1 Data Cleansing

Pada tahapan ini, data dibersihkan dari tanda baca, symbol, angka, dan karakter-karakter lain yang bukan termasuk huruf atau

kata, lalu yang tadinya terdapat huruf kapital pada ulasan akan diubah menjadi huruf kecil seluruhnya. Contoh proses ini dapat dilihat dari perubahan huruf kapital pada kata “Instagram”, “DM”, emoji bintang, tanda baca titik, koma, dan perubahan lainnya.

**Tabel 4.4 Proses *Data Cleansing***

<i>Input</i>	<i>Output</i>
suka banget, bisa jadi tempat curhat dan catatan kecil hidupkuh 😊😊😊	suka banget bisa jadi tempat curhat dan catatan kecil hidupkuh
Sebemer nya aplikasinya bagus simpnhk neko" dan cenderung sedikit bot karena login lewat Instagram. Tapi tolong di tambahkan gitu DM karena jadi gk perlu buka ig. Dan ada bug ketika di alihkan ke ig gk bisa kembali harus di tutup atau gk kembali ke home screen	sebemer nya aplikasinya bagus simpnhk neko dan cenderung sedikit bot karena login lewat instagram tapi tolong di tambahkan gitu dm karena jadi gk perlu buka ig dan ada bug ketika di alihkan ke ig gk bisa kembali harus di tutup atau gk kembali ke home screen
Keren aplikasinya. Coba bisa dipercepat lagi waktu utk posting kontennya pasti lebih seruuuu 🙌🙌	keren aplikasinya coba bisa dipercepat lagi waktu utk posting kontennya pasti lebih seruuuu
Pertama kali log in langsung mengalami bug, tulisan tulisan tidak jelas, tampilan seperti layar handphone rusak, tolong cepat diperbaiki	pertama kali log in langsung mengalami bug tulisan tulisan tidak jelas tampilan seperti layar handphone rusak tolong cepat diperbaiki

### 4.3.2 Normalization

Dalam menulis sebuah ulasan, pengguna cenderung menggunakan Bahasa sehari-hari atau tidak baku, seperti kata

“hidupkuh”, “gk”, “utk”, dan lainnya pada contoh. Sehingga dibutuhkan proses normalisasi bahasa untuk memperbaiki ejaan dan kata-kata yang ditulis tidak baku. Pada tahapan ini, penulis membuat kamus normalisasi secara manual yang menghasilkan sebanyak 1542 kata untuk dinormalisasikan ke kata baku. Contoh proses normalisasi dapat dilihat pada tabel 4.5.

**Tabel 4.5 Proses *Normalization***

<i>Input</i>	<i>Output</i>
suka banget bisa jadi tempat curhat dan catatan kecil <b>hidupkuh</b>	suka banget bisa jadi tempat curhat dan catatan kecil hidup saya
<b>sebemer nya</b> aplikasinya bagus <b>simpnhk</b> <b>neko</b> dan cenderung sedikit bot karena login lewat instagram tapi tolong di tambahkan <b>gitu dm</b> karena jadi <b>gk</b> perlu buka <b>ig</b> dan ada bug ketika di alihkan ke <b>ig gk</b> bisa kembali harus di tutup atau <b>gk</b> kembali ke <b>home screen</b>	sebenarnya aplikasinya bagus simpel tidak aneh dan cenderung sedikit bot karena login lewat instagram tapi tolong di tambahkan begitu pesan singkat karena jadi tidak perlu buka instagram dan ada bug ketika di alihkan ke instagram tidak bisa kembali harus di tutup atau tidak kembali ke beranda screen
keren aplikasinya coba bisa dipercepat lagi waktu <b>utk</b> <b>posting</b> kontennya pasti lebih <b>seruuuu</b>	keren aplikasinya coba bisa dipercepat lagi waktu untuk unggahan kontennya pasti lebih seru
pertama kali <b>log in</b> langsung mengalami bug tulisan tulisan tidak jelas tampilan seperti layar handphone rusak tolong cepat diperbaiki	pertama kali login langsung mengalami bug tulisan tulisan tidak jelas tampilan seperti layar handphone rusak tolong cepat diperbaiki

### 4.3.3 *Stopword Removal*

Tahapan selanjutnya, akan dihilangkan kata-kata yang kerap muncul dalam sebuah ulasan, tetapi tidak memiliki makna yang signifikan atau berpengaruh terhadap ulasan tersebut, seperti kata “banget”, “bisa”, “jadi”, “saya”, “dan”, “ke”, dan yang lainnya. Dalam tahapan ini, penulis memanfaatkan kamus *stopword* yang diciptakan oleh Tala di tahun 2003 pada studinya yang berisi 758 kata. Lalu, penulis membuat kamus *stopword* tambahan secara manual untuk kata-kata yang tidak ada dalam kamus *stopword* yang telah diciptakan oleh Tala dengan tambahan sebanyak 216 kata. Ilustrasi proses *stopwords removal* ini dapat dilihat pada tabel 4.6.

**Tabel 4.6 Proses *Stopword Removal***

<i>Input</i>	<i>Output</i>
suka banget bisa jadi tempat curhat dan catatan kecil hidup saya	suka curhat catatan hidup
sebenarnya aplikasinya bagus simpel tidak aneh dan cenderung sedikit bot karena login lewat instagram tapi tolong di tambahkan begitu pesan singkat karena jadi tidak perlu buka instagram dan ada bug ketika di alihkan ke instagram tidak bisa kembali harus di tutup atau tidak kembali ke beranda screen	aplikasinya bagus simpel aneh cenderung bot login instagram tambahkan pesan singkat buka instagram bug alihkan instagram tutup beranda screen
keren aplikasinya coba bisa dipercepat lagi waktu untuk unggahan kontennya pasti lebih seru	keren aplikasinya coba dipercepat unggahan kontennya seru

<i>Input</i>	<i>Output</i>
pertama kali login langsung mengalami bug tulisan tulisan tidak jelas tampilan seperti layar handphone rusak tolong cepat diperbaiki	login mengalami bug tulisan tulisan tampilan layar handphone rusak tolong cepat diperbaiki

#### 4.3.4 Stemming

Selanjutnya adalah proses *stemming*, yang dilakukan untuk mengubah kata-kata dalam data ulasan menjadi bentuk kata dasarnya, serta menghilangkan imbuhan-imbuhan yang melekat pada kata-kata tersebut, baik imbuhan di awal maupun di akhir. Contohnya dapat dilihat pada tabel 4.7, dimana kata “aplikasinya” menjadi “aplikasi”, “tambahkan” menjadi “tambah”, “dipercepat” menjadi “cepat”, dan lainnya. Untuk tahapan ini, digunakan *library* Sastrawi lagi untuk membantu prosesnya.

**Tabel 4.7 Proses Stemming**

<i>Input</i>	<i>Output</i>
suka curhat catatan hidup	suka curhat catat hidup
aplikasinya bagus simpel aneh cenderung bot login instagram tambahkan pesan singkat buka instagram bug alihkan instagram tutup beranda screen	aplikasi bagus simpel aneh cenderung bot login instagram tambah pesan singkat buka instagram bug alih instagram tutup beranda screen
keren aplikasinya coba dipercepat unggahan kontennya seru	keren aplikasi coba cepat unggah konten seru
login mengalami bug tulisan tulisan tampilan layar handphone rusak tolong cepat diperbaiki	login alami bug tulis tulis tampil layar handphone rusak tolong cepat baik

#### 4.3.4 Tokenizing

Sebelum memasuki tahap pembobotan kata, dilakukan satu Langkah terakhir dari *preprocessing*, yaitu *tokenizing*. Pada tahapan ini, yang tadinya data ulasan berbentuk kalimat yang menyambung, akan dipotong atau dipecah menjadi satuan kata tunggal yang bisa disebut juga dengan token. Berikut adalah hasil perubahan data sebelum dan sesudah proses *tokenizing* pada tabel 4.8.

**Tabel 4.8 Proses *Tokenizing***

<i>Input</i>	<i>Output</i>
suka curhat catat hidup	['suka', 'curhat', 'catat', 'hidup']
aplikasi bagus simpel aneh cenderung bot login instagram tambah pesan singkat buka instagram bug alih instagram tutup beranda screen	['aplikasi', 'bagus', 'simpel', 'aneh', 'cenderung', 'bot', 'login', 'instagram', 'tambah', 'pesan', 'singkat', 'buka', 'instagram', 'bug', 'alih', 'instagram', 'tutup', 'beranda', 'screen']
keren aplikasi coba cepat unggah konten seru	['keren', 'aplikasi', 'coba', 'cepat', 'unggah', 'konten', 'seru']
login alami bug tulis tulis tampil layar handphone rusak tolong cepat baik	['login', 'alami', 'bug', 'tulis', 'tulis', 'tampil', 'layar', 'handphone', 'rusak', 'tolong', 'cepat', 'baik']

#### 4.4 Pembobotan Kata

Setelah data melalui tahapan *preprocessing* dan sudah menjadi lebih bersih untuk diolah ke tahapan selanjutnya, data yang berbentuk kata-kata ini masing-masing akan diberikan nilai atau bobot. Sesuai dengan yang telah dijelaskan di bab sebelumnya, pembobotan ini dinilai penting karena dapat membantu



menentukan frekuensi kemunculan suatu kata dalam dokumen yang bisa berdampak pada kontribusi kata tersebut. Pada penelitian ini, digunakan metode TF-IDF untuk menghitung bobot masing-masing kata. Berikut ilustrasi contoh dokumen atau data ulasan dengan labelnya pada tabel 4.9.

**Tabel 4.9 Dokumen Ulasan**

Dokumen	Data Ulasan	Label
D1	['suka', 'curhat', 'catat', 'hidup']	Positif
D2	['aplikasi', 'bagus', 'simpl', 'aneh', 'cenderung', 'bot', 'login', 'instagram', 'tambah', 'pesan', 'singkat', 'buka', 'instagram', 'bug', 'alih', 'instagram', 'tutup', 'beranda', 'screen']	Negatif
D3	['keren', 'aplikasi', 'coba', 'cepat', 'unggah', 'konten', 'seru']	Positif
D4	['login', 'alami', 'bug', 'tulis', 'tulis', 'tampil', 'layar', 'handphone', 'rusak', 'cepat', 'baik']	Negatif

Perhitungan TF-IDF dimulai dengan menghitung frekuensi kemunculan sebuah kata dalam satu dokumen (TF) sesuai dengan persamaan (4). Berdasarkan tabel 4.9, berikut contoh perhitungan pada dokumen 1:

Hitung *Term Frequency* (TF):

D1:

$$tf_{(D1, 'suka')} = \frac{1}{4} = 0.25$$

$$tf_{(D1, 'curhat')} = \frac{1}{4} = 0.25$$

$$tf_{(D1, 'catat')} = \frac{1}{4} = 0.25$$

$$tf_{(D1, 'hidup')} = \frac{1}{4} = 0.25$$

Lalu dilanjut dengan menghitung kemunculan kata tersebut dalam keseluruhan dokumen (IDF) dengan rumus (5), berikut contoh perhitungannya pada dokumen 1:

Hitung *Inverse Document Frequency* (IDF):

D1:

$$idf_{(D1,suka)} = \log\left(\frac{4}{1}\right) = \log(4) = 0.602$$

$$idf_{(D1,curhat)} = \log\left(\frac{4}{1}\right) = \log(4) = 0.602$$

$$idf_{(D1,catat)} = \log\left(\frac{4}{1}\right) = \log(4) = 0.602$$

$$idf_{(D1,hidup)} = \log\left(\frac{4}{1}\right) = \log(4) = 0.602$$

Kemudian, dari kedua nilai tersebut akan dihitung bobotnya sesuai dengan persamaan TF-IDF (6), berikut contoh perhitungannya pada dokumen 1:

Hitung TF-IDF:

D1:

$$TF - IDF_{(D1,suka)} = tf_{(D1,suka)} \times idf_{(D1,suka)} = 0.25 \times 0.602 = 0.1505$$

$$TF - IDF_{(D1,curhat)} = tf_{(D1,curhat)} \times idf_{(D1,curhat)} = 0.25 \times 0.602 = 0.1505$$

$$TF - IDF_{(D1,catat)} = tf_{(D1,catat)} \times idf_{(D1,catat)} = 0.25 \times 0.602 = 0.1505$$

$$TF - IDF_{(D1,hidup)} = tf_{(D1,hidup)} \times idf_{(D1,hidup)} = 0.25 \times 0.602 = 0.1505$$

Setiap kata pada tabel 4.9, baik pada kolom data ulasan dan kolom label, akan dihitung frekuensinya. Berikut perhitungan untuk D1, D2, D3, dan D4 dijabarkan pada tabel 4.10 dibawah ini:

**Tabel 4.10 Proses Pembobotan Kata**

No.	Term	Term Count				Doc. Count	IDF	TF-IDF			
		D1	D2	D3	D4			D1	D2	D3	D4
1.	Suka	0.25	0	0	0	1	0.602	0.1505	0	0	0
2.	curhat	0.25	0	0	0	1	0.602	0.1505	0	0	0
3.	Catat	0.25	0	0	0	1	0.602	0.1505	0	0	0
4.	Hidup	0.25	0	0	0	1	0.602	0.1505	0	0	0

No.	Term	Term Count				Doc. Count	IDF	TF-IDF			
		D1	D2	D3	D4			D1	D2	D3	D4
5.	aplikasi	0	0.052	0.143	0	2	0.301	0	0.0157	0.043	0
6.	bagus	0	0.052	0	0	1	0.602	0	0.0313	0	0
7.	simpel	0	0.052	0	0	1	0.602	0	0.0313	0	0
8.	aneh	0	0.052	0	0	1	0.602	0	0.0313	0	0
9.	cenderung	0	0.052	0	0	1	0.602	0	0.0313	0	0
10.	bot	0	0.052	0	0	1	0.602	0	0.0313	0	0
11.	login	0	0.052	0	0.09	2	0.301	0	0.0313	0	0.027
12.	Instagram	0	0.158	0	0	1	0.602	0	0.0951	0	0
13.	Tambah	0	0.052	0	0	1	0.602	0	0.0313	0	0
14.	Pesan	0	0.052	0	0	1	0.602	0	0.0313	0	0
15.	Singkat	0	0.052	0	0	1	0.602	0	0.0313	0	0
16.	Buka	0	0.052	0	0	1	0.602	0	0.0313	0	0
17.	Bug	0	0.052	0	0.09	2	0.301	0	0,0157	0	0.027
18.	Alih	0	0.052	0	0	1	0.602	0	0.0313	0	0
19.	Tutup	0	0.052	0	0	1	0.602	0	0.0313	0	0
20.	Beranda	0	0.052	0	0	1	0.602	0	0.0313	0	0
21.	screen	0	0.052	0	0	1	0.602	0	0.0313	0	0
22.	keren	0	0	0.143	0	1	0.602	0	0	0.086	0
23.	Coba	0	0	0.143	0	1	0.602	0	0	0.086	0
24.	Cepat	0	0	0.143	0.09	2	0.301	0	0	0.043	0.027
25.	Unggah	0	0	0.143	0	1	0.602	0	0	0.086	0
26.	Konten	0	0	0.143	0	1	0.602	0	0	0.086	0
27.	seru	0	0	0.143	0	1	0.602	0	0	0.086	0
28.	Alami	0	0	0	0.09	1	0.602	0	0	0	0.0542
29.	Tulis	0	0	0	0.18 2	1	0.602	0	0	0	0.1096
30.	tampil	0	0	0	0.09	1	0.602	0	0	0	0.0542
31.	Layar	0	0	0	0.09	1	0.602	0	0	0	0.0542
32.	Handphone	0	0	0	0.09	1	0.602	0	0	0	0.0542
33.	Rusak	0	0	0	0.09	1	0.602	0	0	0	0.0542
34.	baik	0	0	0	0.09	1	0.602	0	0	0	0.0542

Tabel 4.10 menunjukkan hasil akhir pembobotan TF-IDF berdasarkan empat sampel dokumen. Nilai bobot masing-masing *term* dapat dilihat di empat kolom paling kanan tabel. Pada penelitian ini, pembobotan TF-IDF dilakukan terhadap seluruh dokumen yang berjumlah 3322 dokumen dengan total jumlah *term* sebanyak 4452 *term*.

#### 4.5 Pembagian Data

Sebelum data secara langsung diolah dalam algoritma *Naïve Bayes*, data harus dibagi menjadi dua kategori terlebih dahulu, yaitu data latih (*data training*) dan data uji (*data testing*). Hal ini dilakukan untuk mempersiapkan data sebelum dilakukan proses klasifikasi. Pada penelitian ini, data dibagi menjadi *data training* dan *data testing* dengan proporsi standar yang secara umum digunakan dalam klasifikasi data, yaitu 80% untuk *data training* dan 20% untuk *data testing*. Rincian jumlah *data training* dan *data testing* dapat dilihat pada tabel 4.11.

**Tabel 4.11 Rincian Pembagian Data**

	<b>Positif</b>	<b>Negatif</b>	<b>Total</b>
<b><i>Data Training</i></b>	1235	1422	2657
<b><i>Data Testing</i></b>	+/- 310	+/- 355	665
<b>Total</b>	1548	1774	<b>3322</b>

Pada tabel 4.11 diatas, menunjukkan perbandingan jumlah *data training* dan *data testing* di masing-masing label kelas pada data ulasan pengguna aplikasi *Threads* dengan perbandingan 80:20.

#### 4.6 Pengujian Model *Naïve Bayes*

Pada penelitian ini, *Multinomial Naïve Bayes* digunakan sebagai metode perhitungan untuk melakukan klasifikasi. Proses perhitungan dilakukan baik pada *data training* maupun *data testing* dengan persamaan (1). Setelah seluruh *term* memiliki nilai bobotnya masing-masing melalui pembobotan TF-IDF pada tabel 4.10, akan dimulai perhitungan untuk mencari nilai *prior probability* terlebih dahulu dengan persamaan (2). Prosesnya adalah sebagai berikut:

$$P(c_{positif}) = \frac{2}{4} = 0.5 \qquad P(c_{negatif}) = \frac{2}{4} = 0.5$$

Lalu, perhitungan akan masuk ke tahap pelatihan atau *training*. Berikut adalah perhitungan probabilitas kondisional menggunakan persamaan (3) pada *data training* berdasarkan nilai bobot *term* yang telah dihitung dengan metode TF-IDF pada tabel 4.10. berdasarkan rumus, diketahui bahwa:

$$\text{Nilai } B' = 34$$

$$\text{Total } \textit{term} \textit{ positif} = 13$$

$$\text{Total } \textit{term} \textit{ negatif} = 25$$

$$\begin{aligned} P('suka'|Positif) &= \frac{0.1505 + 1}{13 + 34} \\ &= 0.024 \end{aligned}$$

$$\begin{aligned} P('bug'|Negatif) &= \frac{0.027 + 1}{25 + 34} \\ &= 0.0174 \end{aligned}$$

$$\begin{aligned} P('curhat'|Positif) &= \frac{0.1505 + 1}{13 + 34} \\ &= 0.024 \end{aligned}$$

$$\begin{aligned} P('alih'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('catat'|Positif) &= \frac{0.1505 + 1}{13 + 34} \\ &= 0.024 \end{aligned}$$

$$\begin{aligned} P('tutup'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('hidup'|Positif) &= \frac{0.1505 + 1}{13 + 34} \\ &= 0.024 \end{aligned}$$

$$\begin{aligned} P('beranda'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('aplikasi'|Negatif) &= \frac{0.0157 + 1}{25 + 34} \\ &= 0.017 \end{aligned}$$

$$\begin{aligned} P('screen'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('aplikasi'|Positif) &= \frac{0.043 + 1}{13 + 34} \\ &= 0.022 \end{aligned}$$

$$\begin{aligned} P('keren'|Positif) &= \frac{0.086 + 1}{13 + 34} \\ &= 0.023 \end{aligned}$$

$$\begin{aligned} P('bagus'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('coba'|Positif) &= \frac{0.086 + 1}{13 + 34} \\ &= 0.023 \end{aligned}$$

$$\begin{aligned} P('simpler'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('cepat'|Positif) &= \frac{0.043 + 1}{13 + 34} \\ &= 0.022 \end{aligned}$$

$$\begin{aligned} P('aneh'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\ &= 0.0175 \end{aligned}$$

$$\begin{aligned} P('cepat'|Negatif) &= \frac{0.027 + 1}{25 + 34} \\ &= 0.0174 \end{aligned}$$

$$\begin{aligned}
P('cenderung'|Negatif) &= \frac{0.0313 + 1}{25 + 34} = 0.0175 & P('unggah'|Positif) &= \frac{0.086 + 1}{13 + 34} \\
& & &= 0.023 \\
P('bot'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('konten'|Positif) &= \frac{0.086 + 1}{13 + 34} \\
&= 0.0175 & &= 0.023 \\
P('login'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('seru'|Positif) &= \frac{0.086 + 1}{13 + 34} \\
&= 0.0175 & &= 0.023 \\
P('login'|Negatif) &= \frac{0.027 + 1}{25 + 34} & P('alami'|Negatif) &= \frac{0.0542 + 1}{25 + 34} \\
&= 0.0174 & &= 0.018 \\
P('instagram'|Negatif) &= \frac{0.0951 + 1}{25 + 34} & P('tulis'|Negatif) &= \frac{0.1096 + 1}{25 + 34} \\
&= 0.019 & &= 0.019 \\
P('tambah'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('tampil'|Negatif) &= \frac{0.0542 + 1}{25 + 34} \\
&= 0.0175 & &= 0.018 \\
P('pesan'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('layar'|Negatif) &= \frac{0.0542 + 1}{25 + 34} \\
&= 0.0175 & &= 0.018 \\
P('singkat'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('handphone'|Negatif) &= \frac{0.0542 + 1}{425 + 34} \\
&= 0.0175 & &= 0.018 \\
P('buka'|Negatif) &= \frac{0.0313 + 1}{25 + 34} & P('rusak'|Negatif) &= \frac{0.0542 + 1}{25 + 34} \\
&= 0.0175 & &= 0.018 \\
P('bug'|Negatif) &= \frac{0.0157 + 1}{25 + 34} & P('baik'|Negatif) &= \frac{0.0542 + 1}{25 + 34} \\
&= 0.0172 & &= 0.018
\end{aligned}$$

Setelah proses perhitungan pada 2657 *data training* selesai, maka dilanjutkan untuk menghitung probabilitas kondisional *term* untuk *data testing*-nya. Tahapan ini akan melakukan sebuah *test* pada *data testing* dengan menggunakan *data training* yang telah dihitung pada proses sebelumnya. Dari

665 data yang menjadi *data testing*, akan diambil sampel secara acak pada tabel 4.12 untuk dijadikan contoh pada tahapan ini

**Tabel 4.12 Sampel Data Testing sebelum Preprocessing**

<b>Data Ulasan sebagai Data Testing Sebelum Preprocessing</b>
Masih lebih suka di Twitter, karena bisa menerjemahkan bahasa asing, ini belum bisa dan aneh seharusnya login nya tidak harus pakai Instagram, ada yang ngeluh rusak dan menghapus apk ini sama aja menghapus akun Ig.

Pada tabel 4.12, data ulasan yang akan diuji belum melalui tahap *preprocessing*, maka dari itu pada tabel 4.13 akan dilampirkan data ulasan yang telah bersih dan siap untuk diberikan bobot dan diuji.

**Tabel 4.13 Sampel Data Testing setelah Preprocessing**

<b>Data Ulasan sebagai Data Testing Setelah Preprocessing</b>
['suka', 'twitter', 'terjemah', 'bahasa', 'asing', 'aneh', 'login', 'pakai', 'instagram', 'ngeluh', 'rusak', 'hapus', 'aplikasi', 'hapus', 'akun', 'instagram']

Langkah pertama untuk perhitungan pada *data testing* ini adalah memberikan bobot berdasarkan perhitungan pembobotan yang telah dihitung pada *data training* pada Tabel 4.10. Untuk beberapa *term* yang tidak terdapat pada perhitungan, maka akan diberikan nilai 0 untuk bobotnya, sehingga menghasilkan nilai bobot pada tabel 4.14 dibawah ini:

**Tabel 4.14 Perhitungan TF-IDF untuk Data Testing**

No.	Term	Term Count	IDF	TF-IDF	
		D1		Positif	Negatif
1.	keren	1	0	0	0
2.	twitter	1	0	0	0
3.	terjemah	1	0	0	0
4.	bahasa	1	0	0	0
5.	asing	1	0	0	0
6.	aneh	1	0.602	0	0.0313
7.	login	1	0.301	0	0.0313
8.	pakai	1	0	0	0
9.	instagram	2	0.602	0.0951	0

10.	ngeluh	1	0	0	0
11.	rusak	1	0.602	0	0.0542
12.	hapus	2	0	0	0
13.	aplikasi	1	0.301	0.043	0.0157
14.	akun	1	0	0	0

Setelah selesai diberikan bobot sesuai dengan nilai perhitungan dari *data training*, selanjutnya *data testing* akan masuk ke perhitungan probabilitas kondisional, dengan beberapa nilai yang diketahui:

Nilai  $B' = 34$

Total *term* positif = 13

Total *term* negatif = 25

$$\begin{aligned}
 P('keren'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('keren'|Negatif) &= \frac{0 + 1}{25 + 34} \\
 & & &= 0.0169 \\
 P('twitter'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('twitter'|Negatif) &= \frac{0 + 1}{25 + 34} \\
 & & &= 0.0169 \\
 P('terjemah'|Positif) &= \frac{0 + 1}{13 + 34} & P('terjemah'|Negatif) &= \frac{0 + 1}{25 + 34} \\
 &= 0.021 & &= 0.0169 \\
 P('bahasa'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('bahasa'|Negatif) &= \frac{0 + 1}{25 + 34} \\
 & & &= 0.0169 \\
 P('asing'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('asing'|Negatif) &= \frac{0 + 1}{25 + 34} \\
 & & &= 0.0169 \\
 P('aneh'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('aneh'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\
 & & &= 0.0175 \\
 P('login'|Positif) &= \frac{0 + 1}{13 + 34} = 0.021 & P('login'|Negatif) &= \frac{0.0313 + 1}{25 + 34} \\
 & & &= 0.0175
 \end{aligned}$$



$$P('pakai'|Positif) = \frac{0 + 1}{13 + 34} = 0.021$$

$$P('instagram'|Positif) = \frac{0.0951 + 1}{13 + 34} = 0.023$$

$$P('ngeluh'|Positif) = \frac{0 + 1}{13 + 34} = 0.021$$

$$P('rusak'|Positif) = \frac{0 + 1}{13 + 34} = 0.021$$

$$P('hapus'|Positif) = \frac{0 + 1}{13 + 34} = 0.021$$

$$P('aplikasi'|Positif) = \frac{0.043 + 1}{13 + 34} = 0.022$$

$$P('akun'|Positif) = \frac{0 + 1}{13 + 34} = 0.021$$

$$P('pakai'|Negatif) = \frac{0 + 1}{25 + 34} = 0.0169$$

$$P('instagram'|Negatif) = \frac{0 + 1}{25 + 34} = 0.0169$$

$$P('ngeluh'|Negatif) = \frac{0 + 1}{25 + 34} = 0.0169$$

$$P('rusak'|Negatif) = \frac{0.0542 + 1}{25 + 34} = 0.0179$$

$$P('hapus'|Negatif) = \frac{0 + 1}{25 + 34} = 0.0169$$

$$P('aplikasi'|Negatif) = \frac{0.0157 + 1}{25 + 34} = 0.0172$$

$$P('akun'|Negatif) = \frac{0 + 1}{25 + 34} = 0.0169$$

Berdasarkan perhitungan probabilitas positif dan negatif di atas, tahapan selanjutnya adalah perhitungan dengan rumus *multinomial naïve bayes* pada persamaan (1). Berikut adalah perhitungan untuk probabilitas kelas positif:

$$\begin{aligned} P(\text{positif} | \text{term data testing}) &= P(\text{positif}) \times P('keren' | \text{positif}) \times P('twitter'' | \text{positif}) \\ &\times P('terjemah' | \text{positif}) \times P('bahasa' | \text{positif}) \\ &\times P('asing' | \text{positif}) \times P('aneh' | \text{positif}) \\ &\times P('login' | \text{positif}) \times P('pakai' | \text{positif}) \\ &\times P('instagram' | \text{positif}) \times P('ngeluh' | \text{positif}) \\ &\times P('rusak' | \text{positif}) \times P('hapus' | \text{positif}) \\ &\times P('aplikasi' | \text{positif}) \times P('akun' | \text{positif}) \end{aligned}$$

$$\begin{aligned}
&= 0.5 \times 0.021 \times 0.021 \times 0.021 \times 0.021 \times 0.021 \times 0.021 \\
&\times 0.021 \times 0.021 \times 0.023 \times 0.021 \times 0.021 \times 0.021 \times 0.022 \\
&\times 0.021 = 2,126884983292365912e - 24
\end{aligned}$$

Selanjutnya adalah perhitungan untuk probabilitas kelas negatif:

$$\begin{aligned}
&P(\text{negatif} \mid \text{term data testing}) \\
&= P(\text{negatif}) \times P('suka' \mid \text{negatif}) \\
&\times P('twitter'' \mid \text{negatif}) \times P('terjemah' \mid \text{negatif}) \\
&\times P('bahasa' \mid \text{negatif}) \times P('asing' \mid \text{negatif}) \\
&\times P('aneh' \mid \text{negatif}) \times P('login' \mid \text{negatif}) \\
&\times P('pakai' \mid \text{negatif}) \times P('instagram' \mid \text{negatif}) \\
&\times P('ngeluh' \mid \text{negatif}) \times P('rusak' \mid \text{negatif}) \\
&\times P('hapus' \mid \text{negatif}) \times P('aplikasi' \mid \text{negatif}) \\
&\times P('akun' \mid \text{negatif}) \\
&= 0.5 \times 0.0169 \times 0.0169 \times 0.0169 \times 0.0169 \times 0.0169 \times 0.0175 \\
&\times 0.0175 \times 0.0169 \times 0.0169 \times 0.0169 \times 0.0179 \times 0.0169 \\
&\times 0.0172 \times 0.0169 \\
&= 8,9085254560731825587794839873e - 26
\end{aligned}$$

Berdasarkan hasil perhitungan diatas, dapat dilihat bahwa nilai probabilitas untuk label positif lebih besar dengan nilai 2,126e-24. Maka, dapat disimpulkan bahwa data ulasan yang dijadikan sebagai data uji ini diklasifikasikan sebagai review positif.

#### 4.7 Evaluasi Performa

Setelah melalui proses pemodelan data menggunakan metode *Naïve Bayes*, maka akan dilakukan evaluasi dari model yang telah dihasilkan menggunakan *confusion matrix*. Tujuannya adalah untuk mengevaluasi kinerja model dari proses klasifikasi yang telah dilakukan. Tabel 4.15 berikut melampirkan hasil *confusion matrix*.

**Tabel 4. 15 Hasil *Confusion Matrix* Model *Naïve Bayes***

		<i>Actual Class</i>	
		<b>Positif</b>	<b>Negatif</b>
<i>Predicted Class</i>	<b>Positif</b>	320	55
	<b>Negatif</b>	32	258

Berdasarkan tabel 4.15 diatas, dapat diketahui bahwa nilai *True Positive* (TP) adalah 320, nilai *True Negative* (TN) adalah 258, nilai *False Positive* (FP) adalah 55, dan nilai *False Negative* (FN) adalah 32.

Dari nilai TP, TN, FP, dan FN diatas, maka dapat dihitung nilai *accuracy*, *precision*, *recall*, *specificity*, dan *F-1 score*-nya sesuai dengan persamaan (5), (6), (7), (8), dan (9). Sehingga menghasilkan nilai sebagai berikut:

1. *Accuracy*

$$\frac{320 + 258}{320 + 258 + 32 + 55} = \frac{578}{665} = 0.87$$

2. *Precision*

$$\frac{320}{320 + 55} = \frac{320}{375} = 0.85$$

3. *Recall*

$$\frac{320}{320 + 32} = \frac{320}{352} = 0.91$$

4. *Specificity*

$$\frac{258}{258 + 32} = \frac{258}{290} = 0.90$$

5. *F-1 Score*

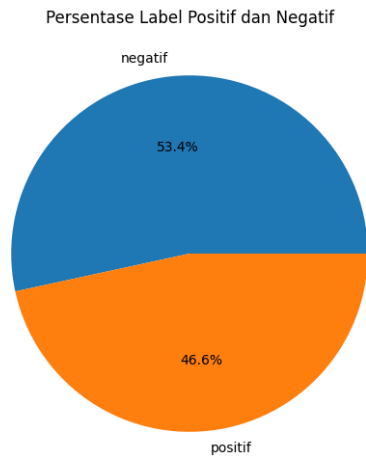
$$\frac{2 * 0.85 * 0.91}{0.85 + 0.91} = \frac{1.547}{1.76} = 0.88$$

Berdasarkan evaluasi yang dilakukan pada pengujian performa model klasifikasi yang dibentuk menggunakan algoritma *Naïve Bayes Classifier*, didapatkan beberapa nilai yang dapat digunakan sebagai tolak ukur seberapa baiknya model yang dihasilkan dalam mengklasifikasikan ulasan yang mengandung sentimen positif dan negative terhadap pengguna aplikasi *Threads*. Evaluasi model pada rasio pembagian data sebesar 80:20 menghasilkan nilai-nilai sebagai berikut:

1. Nilai *accuracy* sebesar 0,87 atau dengan persentase 87% yang menunjukkan nilai hasil rasio prediksi benar dari keseluruhan total data yang diuji
2. Nilai *precision* sebesar 0,85 atau dengan persentase 85% yang menunjukkan rasio dalam memprediksi dengan benar data *true positive* dari keseluruhan data yang diprediksi bernilai positif.
3. Nilai *recall* sebesar 0,91 atau dengan persentase 91% yang menunjukkan rasio dalam memprediksi dengan benar data *true positive* dibandingkan dengan keseluruhan data positif yang sebenarnya.
4. Nilai *specificity* sebesar 0,90 atau dengan persentase 90% yang menunjukkan rasio dalam memprediksi
5. Nilai *f1-score* sebesar 0,88 atau dengan persentase sebesar 88% yang menunjukkan nilai keseimbangan dari *precision* dan *recall*.

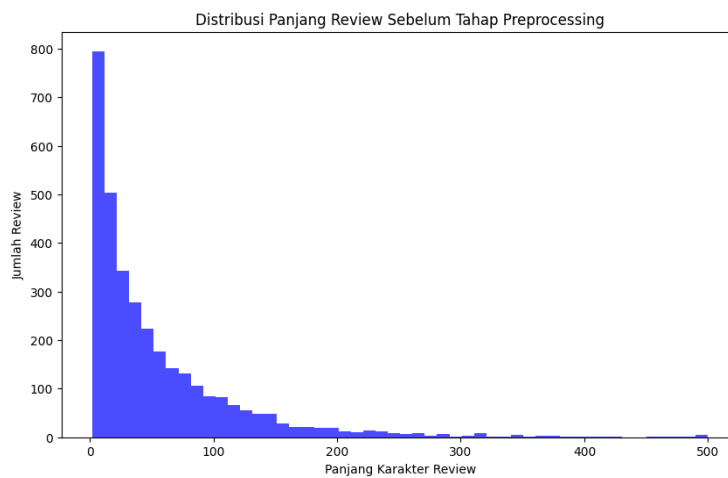
#### **4.8 Analisis Hasil Klasifikasi**

Analisis hasil klasifikasi dilakukan dengan menggambarkan data yang telah diolah, diklasifikasikan, dan dilabel dalam bentuk diagram atau grafik dan *wordcloud* untuk menganalisis informasi yang tersembunyi yang dapat digunakan untuk mengevaluasi kinerja dari *Threads*.

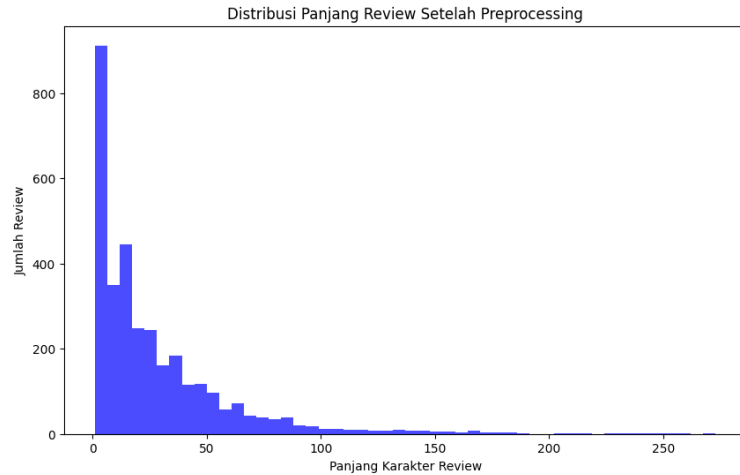


**Gambar 4.2 Diagram Pie Distribusi Sentimen**

Berdasarkan diagram pie diatas, dapat dilihat bahwa mayoritas sentimen dari pengguna aplikasi *Threads* adalah negative, dengan 53.4% ulasan berupa sentimen negatif, sementara sentimen positifnya mencakup 46.6% dari total ulasan.

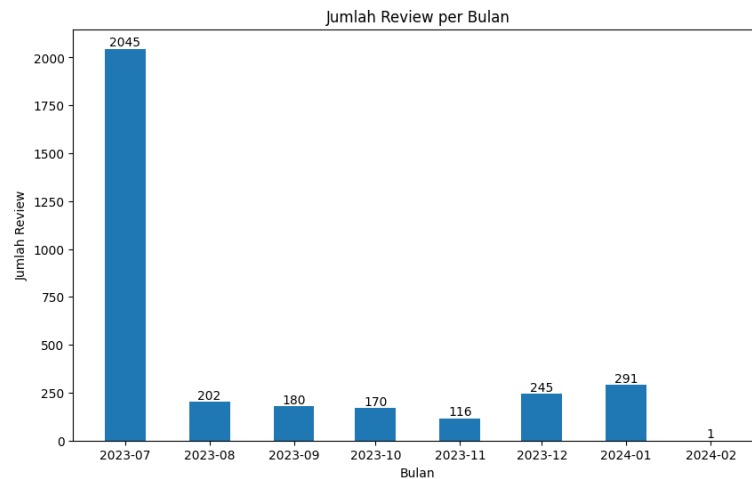


**Gambar 4.3 Grafik Distribusi Panjang Ulasan Sebelum *Preprocessing***



**Gambar 4.4 Grafik Distribusi Panjang Ulasan Setelah *Preprocessing***

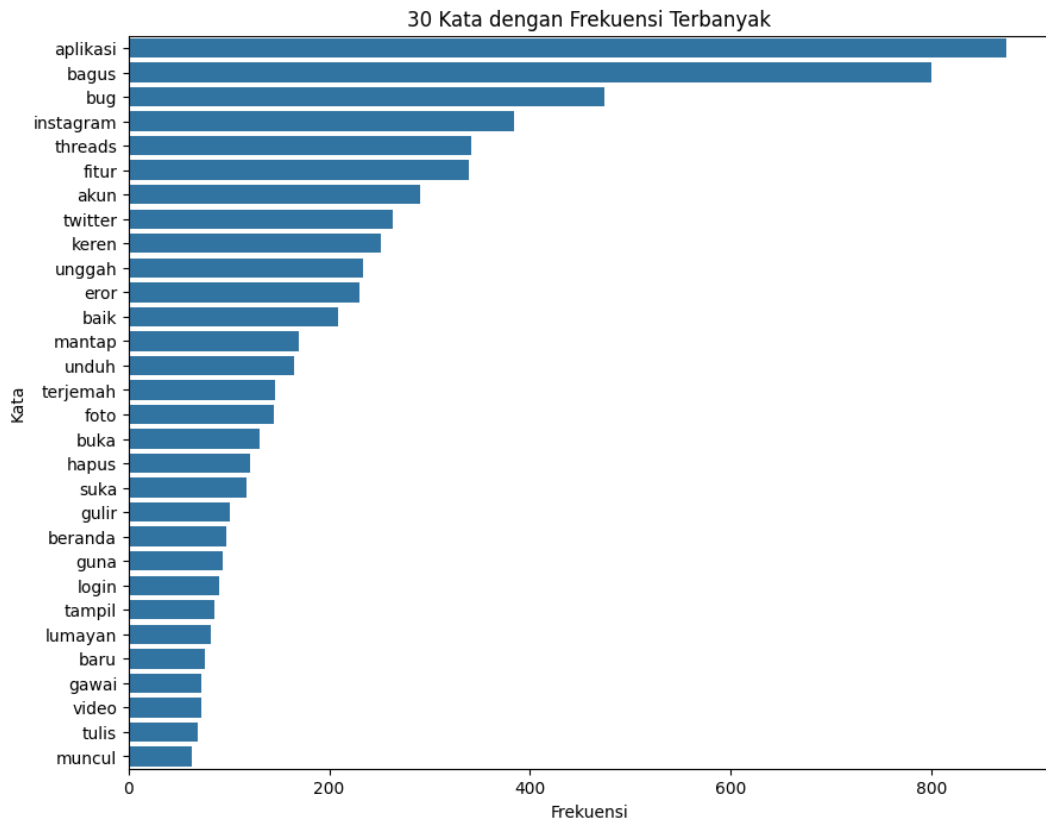
Selanjutnya, dari kedua grafik diatas dapat diperoleh informasi bahwa tahapan *preprocessing* sangat berpengaruh dalam total panjang karakter review. Dari yang sebelumnya bisa mencapai 500 karakter, setelah tahapan *preprocessing* banyak karakter yang berkurang hingga tidak ada yang mencapai 300 karakter, dimana karakter-karakter tersebut mengandung kata-kata yang memiliki relevansi lebih kuat terhadap sentimennya.



**Gambar 4.5 Distribusi Jumlah *Review* per bulan**

Selanjutnya, yaitu total ulasan yang ditulis oleh pengguna per bulannya. Berdasarkan grafiknya, dapat dilihat bahwa mayoritas penggunaannya menuliskan ulasan di bulan Juli 2023, bulan dimana aplikasi baru diluncurkan dan secara langsung menjadi *viral* di kalangan pengguna media sosial, setelah itu di bulan-bulan selanjutnya, jumlah ulasan yang ditulis pun berangsur-angsur semakin

berkurang, hingga pada bulan Desember 2023 dan Januari 2024, mulai sedikit meningkat kembali.



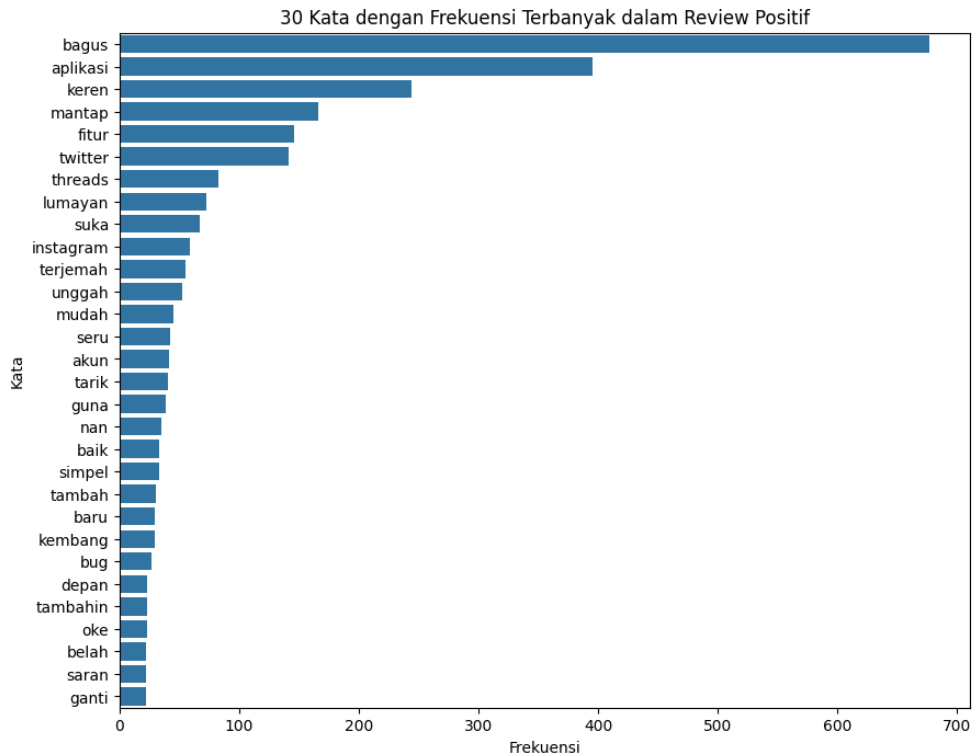
**Gambar 4.6 30 Kata Terbanyak secara Keseluruhan**



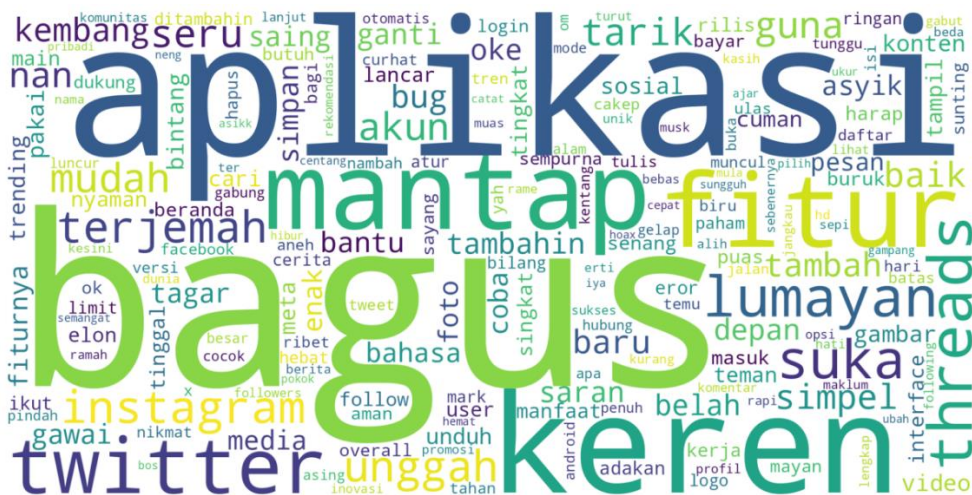
**Gambar 4.7 Word Cloud Keseluruhan Data Ulasan**

Selanjutnya adalah grafik yang menunjukkan 30 kata yang paling banyak muncul secara keseluruhan dan wordcloud-nya. Secara umum, pengguna

memberikan pandangan yang beragam terhadap aplikasi tersebut. Terdapat pujian terhadap fitur-fitur seperti terjemahan, fitur unggah, dan kemudahan penggunaan. Namun, ada juga masukan mengenai bug yang perlu diperbaiki serta permintaan bantuan terkait error atau kesulitan penggunaan.



Gambar 4.8 30 Kata Terbanyak dengan Label Positif



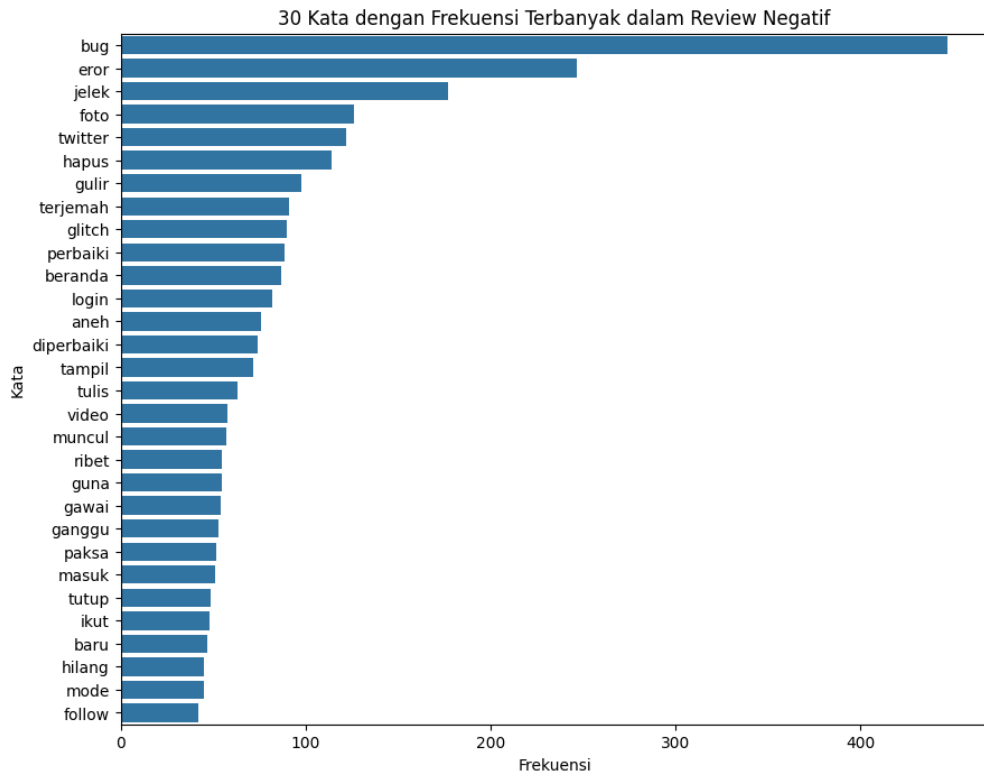
Gambar 4.9 Word Cloud Data Ulasan Positif



Grafik frekuensi 30 kata yang paling banyak muncul serta *word cloud* diatas berguna untuk mengidentifikasi aspek positif dari aplikasi *Threads* yang paling banyak disebutkan oleh pengguna. Seperti kemudahan penggunaan, fitur yang bagus, serta bagaimana pengalaman pengguna yang baik saat memakai aplikasi *Threads*. Kata “bagus” muncul paling banyak, lalu diikuti dengan “aplikasi”, “keren”, “mantap”, “fitur”, dan kata-kata lainnya dalam ulasan yang bersentimen positif dari pengguna. Dari grafik dan *wordcloud* ulasan pengguna yang bersentimen positif tersebut dapat ditarik beberapa kesimpulan sebagai berikut:

1. Kualitas dan Kinerja Aplikasi: Kata-kata seperti "bagus", "keren", "mantap", "fitur", dan "mudah" menunjukkan bahwa pengguna merasa puas dengan kualitas dan kinerja aplikasi *Threads*, serta fitur-fitur yang tersedia pada aplikasi.
2. Pengalaman Pengguna yang Positif: Kata-kata seperti "seru", "lumayan", "suka", dan "tarik" menunjukkan bahwa pengguna memiliki pengalaman yang baik saat menggunakan aplikasi ini, dengan antarmuka yang menarik dan memberikan kemudahan dalam penggunaannya.
3. Kemudahan dan Kesederhanaan: Kata-kata seperti "simplen" dan "mudah" menunjukkan bahwa pengguna menyukai kesederhanaan dan kemudahan penggunaan aplikasi *Threads*.
4. Harapan untuk Perbaikan: Meskipun ada kata "bug" yang muncul, hal ini dapat diartikan bahwa pengguna mengharapkan perbaikan dan pembaruan dari pengembang aplikasi, namun hal ini tidak secara signifikan mengurangi sentimen positif mereka terhadap aplikasi.

Dengan demikian, dapat disimpulkan bahwa pengguna aplikasi *Threads* yang menuliskan ulasan dengan sentimen positif menganggap aplikasi ini memiliki kualitas, kinerja, dan fungsionalitas yang baik, memberikan pengalaman pengguna yang memuaskan, mudah digunakan, serta memberikan manfaat yang signifikan dalam penggunaan sehari-hari mereka.



**Gambar 4.10 30 Kata Terbanyak dengan Label Negatif**



**Gambar 4.11 Word Cloud Data Ulasan Negatif**

Selanjutnya adalah grafik dan *word cloud* yang berguna untuk mengidentifikasi aspek negatif dari aplikasi *Threads* yang paling banyak disebutkan oleh pengguna. Seperti kekurangan fitur dalam aplikasi, tantangan dalam penggunaan, hingga masalah teknis yang terjadi saat memakai aplikasi *Threads*. Kata “*bug*” muncul paling banyak, lalu diikuti dengan “*eror*”, “*jelek*”, “*foto*”, “*twitter*”, dan kata-kata lainnya dalam ulasan yang bersentimen negatif

dari pengguna. Berdasarkan grafik dan *wordcloud* ulasan pengguna yang bersentimen negatif tersebut dapat ditarik beberapa kesimpulan sebagai berikut:

1. Masalah Teknis dan Bug: Kata-kata seperti "bug", "eror", dan "hapus" menunjukkan adanya masalah teknis, bug, atau kesalahan dalam penggunaan aplikasi yang mengganggu pengalaman pengguna.
2. Kualitas Konten dan Interaksi: Kata-kata seperti "instagram", "twitter", "foto", "video", dan "terjemah" menunjukkan bahwa pengguna mengalami masalah terkait kualitas konten, pengalaman berbagi foto/video, terjemahan konten, atau kemampuan untuk menulis, ataupun berinteraksi dengan pengguna lain.
3. Ketidakpuasan dengan Fungsionalitas: Kata-kata seperti "jelek", "ribet", "perbaiki", dan "paksa" menunjukkan bahwa pengguna secara langsung mengekspresikan ketidakpuasannya dengan fungsionalitas, performa, atau pengalaman penggunaan secara keseluruhan.

Dengan demikian, berdasarkan simpulan dari hasil analisis sentimen positif dan negatif diatas, berikut beberapa rekomendasi yang dapat diberikan untuk meningkatkan kinerja dan kualitas aplikasi Threads:

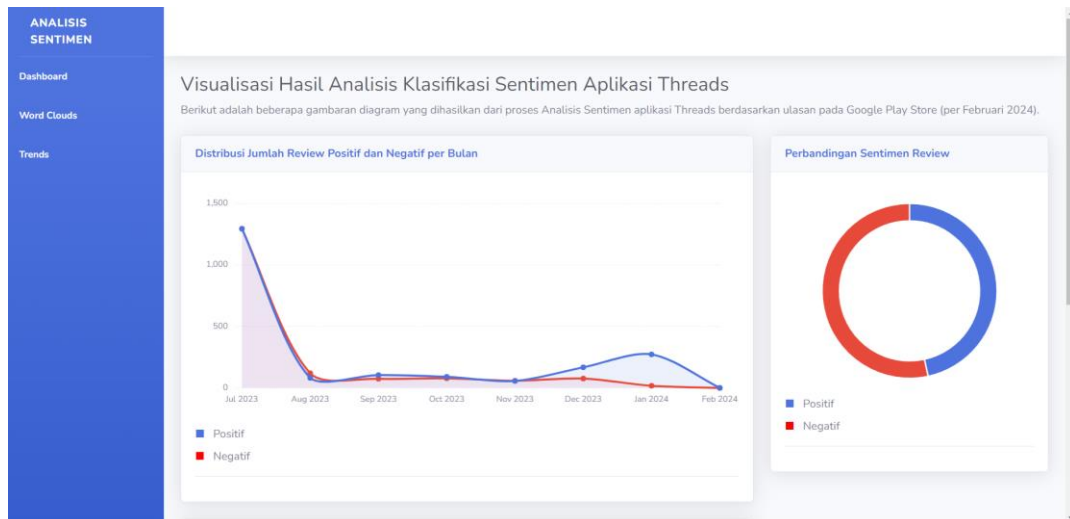
1. Perbaiki teknis dan *bug*:
  - a. Melakukan pemantauan rutin terhadap kemungkinan *bug* dan masalah teknis dalam aplikasi.
  - b. Mengimplementasikan mekanisme yang memungkinkan pengguna untuk melaporkan *bug* atau masalah secara langsung dan menyediakan solusi yang cepat dan efisien.
  - c. Melakukan optimasi performa aplikasi untuk memastikan kinerja yang lancar dan responsif, terutama saat membuka halaman beranda atau mengunggah konten.
2. Peningkatan fungsionalitas dan pengalaman pengguna:
  - a. Menambahkan fitur-fitur baru yang dapat meningkatkan pengalaman pengguna dalam mengoperasikan aplikasi.
  - b. Memperbarui tampilan, sehingga menjadi lebih menarik dan mudah dipahami oleh pengguna baru.

3. Optimalisasi konten dan interaksi:
  - a. Meningkatkan kualitas konten yang disajikan, seperti foto, video, dan teks, sehingga lebih menarik dan relevan bagi pengguna.
  - b. Menyediakan fitur-fitur yang memungkinkan pengguna untuk berinteraksi dengan konten dan pengguna lain dengan lebih efektif dan menyenangkan.
4. Meningkatkan Layanan dan Dukungan Pengguna:
  - a. Menyediakan layanan pelanggan yang responsif dan informatif, termasuk tutorial penggunaan, panduan *troubleshooting*, dan bantuan langsung dalam mengatasi masalah.
  - b. Meningkatkan komunikasi dengan pengguna melalui umpan balik, survei kepuasan, atau forum pengguna untuk memahami kebutuhan dan harapan pengguna secara lebih baik.

Dengan menerapkan rekomendasi-rekomendasi ini, diharapkan aplikasi *Threads* dapat meningkatkan kinerja, kualitas, dan pengalaman pengguna secara keseluruhan, sehingga lebih memuaskan dan relevan bagi pengguna.

#### **4.9 Visualisasi**

Setelah melewati serangkaian tahapan, tahapan terakhir adalah menerapkan visualisasi untuk menampilkan hasil analisis data. Proses visualisasi dilakukan melalui pembuatan *dashboard* berbasis web menggunakan Bahasa pemrograman HTML, CSS, dan JS. Visualisasi ini dapat membantu dalam pemahaman pola, tren, dan *insight* yang terdapat dalam data yang telah dianalisis secara lebih mudah. Berikut adalah tampilan visualisasi berbasis web yang telah dibuat:



**Gambar 4.12 Halaman *Dashboard Web* (1)**



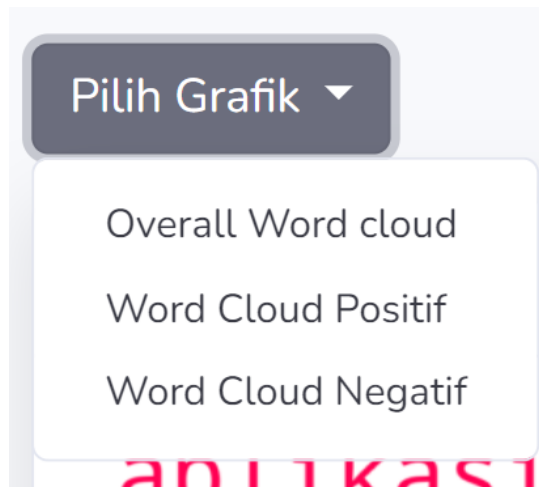
**Gambar 4.13 Halaman *Dashboard Web* (2)**

Gambar 4.12 dan 4.13 diatas merupakan tampilan dari halaman *dashboard* yang menjadi halaman utama dalam web yang dibuat. Halaman ini menampilkan beberapa visualisasi hasil analisis klasifikasi dengan menggunakan grafik berupa diagram batang, *pie chart*, dan diagram garis.



**Gambar 4.14** Halaman *Word Cloud Web*

Selanjutnya adalah halaman dari *word cloud* yang menampilkan visualisasi dari data *review* yang ditulis oleh pengguna pada aplikasi *Threads*. Pada halaman ini terdapat fitur *dropdown* yang berisi 3 pilihan, yaitu *word cloud* secara keseluruhan, *word cloud* dengan sentimen positif, dan *word cloud* dengan sentiment negatif.

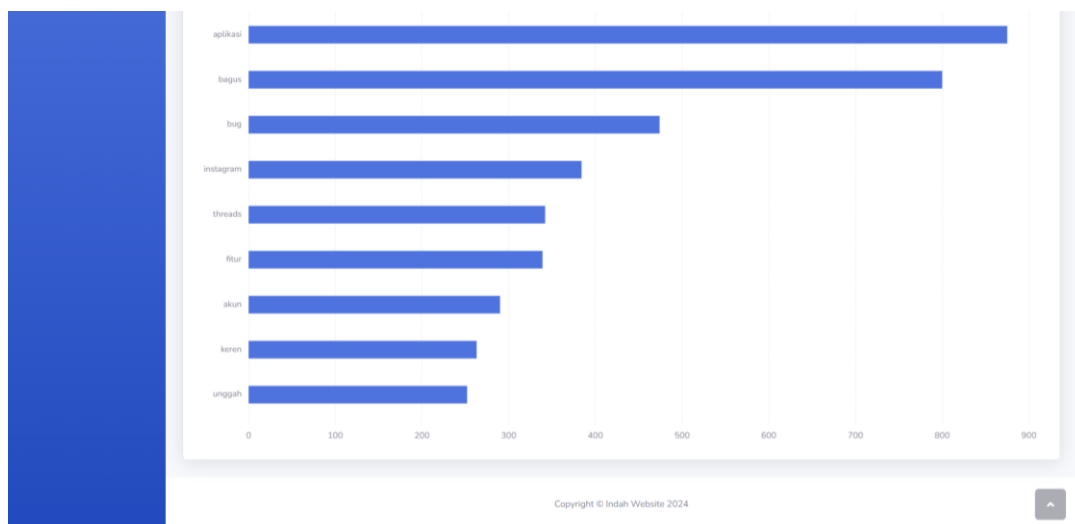


**Gambar 4.15** Fitur *Dropdown* halaman *Word Cloud*

Dengan adanya fitur *dropdown* ini, dapat memudahkan pengguna untuk melihat apa kata-kata yang paling banyak disebutkan ketika pengguna aplikasi *Threads* menuliskan sebuah ulasan.

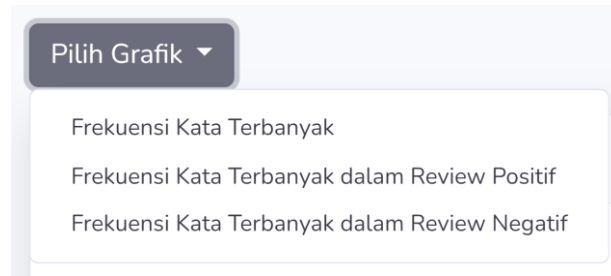


**Gambar 4.16 Halaman Trends Web (1)**



**Gambar 4. 17 Halaman Trends Web (2)**

Halaman terakhir dari web ini adalah halaman *trends*. Halaman ini berisi diagram batang dengan 10 kata yang paling banyak ditulis dalam ulasan aplikasi *Threads* di *Google Play Store*. Dengan representasi diagram yang jelas, pengguna akan lebih mudah untuk melihat tren dan pola terkait penggunaan aplikasi *Threads* ini. Sama dengan halaman *word cloud*, pada halaman *trends* juga disediakan fitur *dropdown* dengan tiga pilihan, yaitu dari data secara keseluruhan, data bersentimen positif, dan data yang bersentimen negatif.



**Gambar 4.18** Fitur *Dropdown* Halaman *Trends*

Fitur *dropdown* ini dapat memudahkan pengguna dan memberikan pengalaman yang menarik untuk menganalisis dan memahami bagaimana tren penggunaan aplikasi *Threads* berdasarkan kata-kata yang banyak muncul pada data. Setelah visualisasi selesai dibuat, maka proses analisis sentimen dapat dinyatakan selesai.