

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.

## BAB 4

### ANALISA DAN HASIL

#### 4.1 Pengumpulan Data

Penelitian Tugas Akhir ini menggunakan data ulasan pengguna aplikasi Bank Jago dari ulasan *Google Play Store*. Berdasarkan data yang ada pada *play store*, aplikasi Bank Jago mulai dirilis pada tanggal 13 April 2021. Pada saat ini bank jago telah diunduh lebih dari 10 juta kali dan memiliki sebanyak 161 ribuan review pada *Google Play Store*. Penggunaan jumlah data yang diambil ialah sebanyak 1000 data terbaru yang ada pada ulasan aplikasi tersebut. Pengambilan data dilakukan dengan cara melakukan *scraping data* menggunakan *tools google co-laboratory* dan Bahasa pemrograman *Python*. Berikut adalah tampilan modul dari pengambilan data yang dilakukan pada Gambar 4.1.

```
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.jago.digitalBanking',
    lang='id',
    country='id',
    sort=Sort.NEWEST,
    count=1000,
    filter_score_with=None
)
```

**Gambar 4.1.** Modul *Scraping*

. Berdasarkan Gambar 4.1, *scraping data* pada ulasan aplikasi dilakukan menggunakan API *google play scraper*. Data yang diambil keseluruhannya berbahasa Indonesia dengan jumlah 1.000 data. Contoh hasil dari *scraping data* yang telah dilakukan tersebut dapat dilihat pada Tabel 4.1

**Tabel 4.1.** Contoh *Scraping Data*

No	Username	Date	Content
1	Oid Dagneel	2024-07-05	Alhamdulillah
2	Yanto Sanggel	2024-07-05	gimana sih akun saya di mintai verifikasi terus pdahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah gk bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini

**Hak Cipta Dilindungi Undang-Undang**

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

**Tabel 4.1.** Contoh *Scraping Data* (Tabel Lanjutan...)

No	Username	Date	Content
3	Miku Can	2024-07-05	Membantu bangetttt, makasih.
4	capekkdighosting	2024-07-05	Cs nya lambat. Minta solusi jawabanya tanya@jago terus.. Berulang kali di kasih jawaban kaya gitu.
5	Wildan Jauharuddin	2024-07-05	Bagus dan sangat membantu, bank yang sangat dapat dipercaya

## 4.2 Preprocessing

*Preprocessing* adalah proses awal pada tahap *text mining* dimana mengubah data menjadi lebih terstruktur. Pada tahap ini, data ulasan yang sudah didapatkan dari proses *web scraping* akan dilakukan *preprocessing*. Proses pada *preprocessing* ini meliputi *cleaning*, *case folding*, normalisasi, *stopwords*, *tokenizing*, dan *stemming*.

### 4.2.1 Cleaning

Fase awal dalam *preprocessing* yaitu melakukan *cleaning*, yaitu membersihkan atribut pada teks yang tidak terdapat hubungannya dengan informasi ataupun yang tidak dibutuhkan seperti menghapus emoji, karakter dan termasuk didalamnya *case folding* atau melakukan perubahan pada semua huruf menjadi huruf kecil. Di bawah ini merupakan tampilan modul dari *cleaning* dapat dilihat pada Gambar 4.2.

```
# Cleaning text
from cleantext import clean

# Menghapus Emoji
def remove_emoji(text):
    text = clean(text, no_emoji=True)
    return text

# Menghapus karakter
def clean_text(text):
    text = re.sub(r"\n|w+:@\w+|(@\w+|(\[A-Za-z0-9+\])\d+|([\^0-9A-Za-z \t])|(\w+:\w+\S+)|^rt|http.+?'", ' ', text)
    return text

# Mengubah menjadi huruf kecil
def lower(text):
    return text.lower()
```

**Gambar 4.2.** Modul *Cleaning*

Setelah dijalkannya modul pada Gambar 4.2, maka didapatkan hasil dari *cleaning*. Berikut ini adalah contoh hasil dari *cleaning* yang dapat dilihat pada Tabel 4.2.



Hak Cipta Dilindungi Undang-Undang

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

Tabel 4.2. Contoh *Cleaning*

No.	Sebelum <i>Cleaning</i>	Setelah <i>Cleaning</i>
	Alhamdulillah gimana sih akun saya di mintai verifikasi terus pdahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah gk bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini Membantu bangetttt, makasih Cs nya lambat. Minta solusi jawabanya tanya@jago terus.. Berulang kali di kasih jawaban kaya gitu? Bagus dan sangat membantu, bank yang sangat dapat dipercaya	alhamdulillah gimana sih akun saya di mintai verifikasi terus pdahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah gk bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini Membantu bangetttt makasih cs nya lambat minta solusi jawabanya tanya terus berulang kali di kasih jawaban kaya gitu bagus dan sangat membantu bank yang sangat dapat dipercaya

4.2.2 Normalisasi

Setelah melakukan *cleaning*, selanjutnya akan dilakukan proses normalisasi yaitu mengembalikan kata tidak baku menjadi baku. Pada normalisasi dilakukan penggunaan kamus slang. Kamus slang yang digunakan adalah *colloquial lexicon*. Berikut ini adalah tampilan modul dari proses normalisasi yang dapat dilihat pada Gambar 4.3.

```
# Normalisasi
slang_dict = pd.Series(kamus_slang['formal'].values, index=kamus_slang['slang']).to_dict()

def replace_slang(text, slang_dict):
    words = text.split()
    replaced_text = ' '.join([slang_dict.get(word, word) for word in words])
    return replaced_text
```

Gambar 4.3. Modul Normalisasi

Setelah dijalankannya modul yang ada pada Gambar 4.3, maka akan didapatkan hasil dari normalisasi. Berikut ini adalah contoh hasil dari normalisasi yang telah dilakukan yang dilihat pada Tabel 4.3.

Tabel 4.3. Contoh Normalisasi

No.	Sebelum Normalisasi	Setelah Normalisasi
	alhamdulillah	alhamdulillah

**Hak Cipta Dilindungi Undang-Undang**

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

**Tabel 4.3.** Contoh Normalisasi (Tabel Lanjutan...)

No.	Sebelum Normalisasi	Setelah Normalisasi
	gimana sih akun saya di mintai verifikasi terus padahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah gk bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini Membantu bangettttt makasih cs nya lambat minta solusi jawabanya tanya terus berulang kali di kasih jawaban kaya gitu bagus dan sangat membantu bank yang sangat dapat dipercaya	bagaimana sih akun saya di mintai verifikasi terus padahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah enggak bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini membantu banget terima kasih cs nya lambat meminta solusi jawabanya tanya terus berulang kali di kasih jawaban kayak begitu bagus dan sangat membantu bank yang sangat dapat dipercaya

**4.2.3 Tokenizing**

Setelah dilakukan proses normalisasi, selanjutnya dilakukan proses *tokenizing* yaitu mengubah dokumen menjadi atribut perkata. Berikut adalah tampilan modul dari proses *tokenizing* yang dapat dilihat pada Gambar 4.4.

```
# Tokenizing
def tokenize(d):
    return word_tokenize(d)

# Implementasi tokenizing
tokenizing = formal.copy()
tokenizing['Text_Tokenize'] = tokenizing['Text_Formal'].apply(tokenize)
tokenizing.drop(columns='Text_Formal', inplace=True)
tokenizing.head()
```

**Gambar 4.4.** Modul *Tokenizing*

Setelah dijalankannya modul pada Gambar 4.4, maka didapatkan hasil dari proses *tokenizing*. Berikut ini adalah contoh hasil dari *tokenizing* yang dapat dilihat pada Tabel 4.4.

**Tabel 4.4.** Contoh *Tokenizing*

No.	Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
	alhamdulillah	['alhamdulillah']

**Tabel 4.4.** Contoh *Tokenizing* (Tabel Lanjutan...)

No.	Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
5	<p>bagaimana sih akun saya di mintai verifikasi terus padahal itu akun saya sendiri pas buka akun di arahin ganti pasword setelah ganti pasword malah enggak bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini</p> <p>membantu banget terima kasih</p> <p>cs nya lambat meminta solusi jawabanya tanya terus berulang kali di kasih jawaban kayak begitu</p> <p>bagus dan sangat membantu bank yang sangat dapat dipercaya</p>	<p>['bagaimana', 'sih', 'akun', 'saya', 'di', 'mintai', 'verifikasi', 'terus', 'padahal', 'itu', 'akun', 'saya', 'sendiri', 'pas', 'buka', 'akun', 'di', 'arahin', 'ganti', 'password', 'setelah', 'ganti', 'password', 'malah', 'enggak', 'bisa', 'masuk', 'dalam', 'akun', 'terus', 'dalam', 'akun', 'itu', 'bukan', 'kosong', 'tapi', 'ada', 'saldo', 'nya', 'parah', 'sekali', 'admin', 'dari', 'pihak', 'bank', 'jago', 'digital', 'ini']</p> <p>['membantu', 'banget', 'terima', 'kasih']</p> <p>['cs', 'nya', 'lambat', 'meminta', 'solusi', 'jawabanya', 'tanya', 'terus', 'berulang', 'kali', 'di', 'kasih', 'jawaban', 'kayak', 'begitu']</p> <p>['bagus', 'dan', 'sangat', 'membantu', 'bank', 'yang', 'sangat', 'dapat', 'dipercaya']</p>

#### 4.2.4 *Remove Stopword*

Setelah melakukan *tokenizing*, proses selanjutnya yang dilakukan adalah *stopword*, yaitu mengurangi jumlah kata dalam dokumen. Berikut adalah tampilan modul dari proses *stopword* yang dapat dilihat pada Gambar 4.5.

```
# Remove stopword/Filtering
def remove_stopwords(text):
    # Muat stopwords default untuk bahasa Indonesia
    stop_words = set(stopwords.words('indonesian'))

    # Hapus kata-kata penting yang tidak ingin dihapus
    custom_stop_words = stop_words - {'tidak', 'dapat', 'guna', 'jika', 'sementara', 'belum', 'tetapi', 'kecuali',
                                     'tapi', 'ada', 'tanpa', 'ok', 'hanya', 'kurang', 'seperti', 'sepertinya', 'harus', 'harusnya',
                                     'tolong', 'bisa', 'sangat', 'paling', 'semakin'}

    # Jika input berupa list, gabungkan menjadi satu string
    if isinstance(text, list):
        text = ' '.join(text)

    # Filter teks dengan stopwords yang sudah disesuaikan
    filtered_words = [word for word in text.split() if word.lower() not in custom_stop_words]

    return filtered_words
```

**Gambar 4.5.** Modul *Stopword*

Setelah dijalkannya modul pada Gambar 4.5, maka akan didapatkan

- 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.

**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.

hasil dari *stopword*. Berikut ini adalah hasil contoh dari *Stopword* yang dapat dilihat pada Tabel 4.5.

**Tabel 4.5.** Contoh *Stopword*

No.	Sebelum <i>Stopword</i>	Setelah <i>Stopword</i>
	alhamdulillah	['alhamdulillah']
	bagaimana sih akun saya di mintai verifikasi terus padahal itu akun saya sendiri pas buka akun di arahin ganti password setelah ganti password malah enggak bisa masuk dalam akun terus dalam akun itu bukan kosong tapi ada saldo nya parah sekali admin dari pihak bank jago digital ini membantu banget terima kasih	['sih', 'akun', 'mintai', 'verifikasi', 'akun', 'pas', 'buka', 'akun', 'arahin', 'ganti', 'password', 'ganti', 'password', 'bisa', 'masuk', 'akun', 'akun', 'kosong', 'tapi', 'ada', 'saldo', 'nya', 'parah', 'admin', 'bank', 'jago', 'digital']
	cs nya lambat meminta solusi jawabanya tanya terus berulang kali di kasih jawaban kayak begitu	['membantu', 'banget', 'terima', 'kasih'] ['cs', 'nya', 'lambat', 'solusi', 'jawabanya', 'berulang', 'kali', 'kasih', 'kayak']
5	bagus dan sangat membantu bank yang sangat dapat dipercaya	['bagus', 'sangat', 'membantu', 'bank', 'sangat', 'dapat', 'dipercaya']

#### 4.2.5 Stemming

Setelah melakukan *stopword*, selanjutnya dilakukan proses terakhir dari *pre-processing* yaitu *stemming*. *Stemming* adalah proses merubah sebuah kata imbuhan menjadi kata dasar. Berikut adalah tampilan modul dari proses *stemming* pada Gambar 4.6.

```
# Stemming
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

factory = StemmerFactory()
stemmer = factory.create_stemmer()
ps = nltk.PorterStemmer()

def stemming(d):
    return [stemmer.stem(word) for word in d]
```

**Gambar 4.6.** Modul *Stemming*

Setelah dijalkannya modul pada Gambar 4.6, maka akan didapatkan hasil dari *stemming*. Berikut ini adalah contoh hasil dari proses *stemming* yang telah dilakukan dapat dilihat pada Tabel 4.6.



**Hak Cipta Dilindungi Undang-Undang**

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

**Tabel 4.6.** Contoh *Stemming*

No.	Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
	'alhamdulillah'	['alhamdulillah']
	'sih', 'akun', 'mintai', 'verifikasi', 'akun', 'pas', 'buka', 'akun', 'arahin', 'ganti', 'password', 'ganti', 'password', 'bisa', 'masuk', 'akun', 'akun', 'kosong', 'tapi', 'ada', 'saldo', 'nya', 'parah', 'admin', 'bank', 'jago', 'digital'	['sih', 'akun', 'minta', 'verifikasi', 'akun', 'pas', 'buka', 'akun', 'arah', 'ganti', 'password', 'ganti', 'password', 'bisa', 'masuk', 'akun', 'akun', 'kosong', 'tapi', 'ada', 'saldo', 'nya', 'parah', 'admin', 'bank', 'jago', 'digital']
	'membantu', 'banget', 'terima', 'kasih'	['bantu', 'banget', 'terima', 'kasih']
	'cs', 'nya', 'lambat', 'solusi', 'jawabanya', 'berulang', 'kali', 'kasih', 'kayak'	['cs', 'nya', 'lambat', 'solusi', 'jawabanya', 'ulang', 'kali', 'kasih', 'kayak']
	'bagus', 'sangat', 'membantu', 'bank', 'sangat', 'dapat', 'dipercaya'	['bagus', 'sangat', 'bantu', 'bank', 'sangat', 'dapat', 'percaya']

Setelah dilakukannya semua proses *preprocessing*, maka selanjutnya kita melakukan penyusunan kata-kata tersebut kembali menjadi sebuah kalimat. Berikut ini adalah contoh hasil dari seluruh proses *preprocessing* yang kembali disusun menjadi sebuah kalimat yang dapat dilihat pada Tabel 4.7.

**Tabel 4.7.** Contoh Hasil *Preprocessing*

No.	Hasil <i>Preprocessing</i>
1	alhamdulillah
2	sih akun minta verifikasi akun pas buka akun arahin ganti password ganti password bisa masuk akun akun kosong tapi ada saldo nya parah admin bank jago digital
3	bantu banget terima kasih
4	cs nya lambat solusi jawabanya ulang kali kasih kayak
5	bagus sangat bantu bank sangat dapat percaya

**4.3 Pelabelan Data**

Setelah dilakukannya proses *preprocessing*, selanjutnya dilakukan proses yang disebut dengan pelabelan data. Pada penelitian ini dilakukan pelabelan dengan menjadikan rating yang diberikan oleh pengguna sebagai pedoman dan dibagi menjadi kategori positif dan negatif. Didapatkan hasil pelabelan positif sebanyak 44 data dan negatif sebanyak 305 data. Di bawah adalah beberapa contoh dari pelabelan data pada Tabel 4.8.



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

Tabel 4.8. Contoh Hasil Pelabelan Data

Ulasan	Sentimen
alhamdulillah	positif
sih akun minta verifikasi akun pas buka akun	negatif
arahin ganti password ganti password bisa masuk	
akun akun kosong tapi ada saldo nya parah admin	
bank jago digital	
bantu banget terima kasih	positif
cs nya lambat solusi jawabanya ulang kali kasih	negatif
kayak	
bagus sangat bantu bank sangat dapat percaya	positif

4.4 Pembobotan TF-IDF

Pembobotan TF-IDF berguna dalam melakukan pembobotan terhadap ke-munculan suau data terhadap sebuah dokumen. *Term Frequency* (TF) menampilkan jumlah kemunculan suatu kata dalam dokumen sedangkan *Inverse Document Frequency* (IDF) menunjukkan tingkat kepentingan suatu kata dalam kumpulan dokumen. Berikut ini adalah hasil dari pembobotan TF-IDF yang dapat dilihat pada Gambar 4.7.

	ada	admin	aju	akses	aktif	akun	alami	alhamdulillah	aman	ampun	...	ulang
0	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	1.000000	0.0	0.0	...	0.0
1	0.116056	0.187642	0.0	0.0	0.000000	0.716014	0.0	0.000000	0.0	0.0	...	0.0
2	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.419727	0.0	0.0	...	0.0
3	0.000000	0.000000	0.0	0.0	0.59308	0.000000	0.0	0.000000	0.0	0.0	...	0.0
4	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...
954	0.150842	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0
955	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0
956	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0
957	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0
958	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.0	0.0	...	0.0

Gambar 4.7. Pembobotan TF-IDF

4.5 Hasil Klasifikasi

Selanjutnya data yang telah didapat akan dilakukan klasifikasi menggunakan algoritma *Support Vector Machine*(SVM). Pengklasifikasian dilakukan dengan menggunakan kernel linier dengan nilai  $c=1$ . Pengklasifikasian dibagi menjadi 3 skenario yang menghasilkan nilai akurasi. Skenario pertama yaitu terdiri dari 80% data *training* dan 20% data *testing*, skenario kedua terdiri dari 70% data *training* dan 30% data *testing*, dan skenario ketiga terdiri dari 60% data *training* dan



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.

40% data *testing*. Berikut ini merupakan hasil pengujian dari algoritma SVM pada Tabel 4.9.

**Tabel 4.9.** Hasil Pengujian klasifikasi SVM

Skenario	Kernel	Accuracy
80:20	Linear	0,88
70:30	Linear	0,87
60:40	Linear	0,88

Berikut ini adalah report hasil klasifikasi SVM yang telah dilakukan

1. Skenario Pertama 80:20

Skenario yang pertama adalah skenario 80:20 dengan data *testing* sebesar 20% dengan jumlah data *testing* sebanyak 192 data dan data *training* sebanyak 767 data. Menghasilkan nilai *accuracy* sebesar 0,8802 seperti yang terlihat pada Gambar 4.8.

```
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy:2f}')

Accuracy: 0.880208
```

**Gambar 4.8.** Klasifikasi Skenario Pertama

2. Skenario Kedua 70:30

Selanjutnya klasifikasi skenario kedua 70:30 dengan data *testing* sebesar 30% digunakan jumlah data *testing* sebanyak 288 data dan data *training* sebanyak 671 data menghasilkan nilai *accuracy* sebesar 0,875 dapat dilihat pada Gambar 4.9.

```
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy:2f}')

Accuracy: 0.875000
```

**Gambar 4.9.** Klasifikasi Skenario Kedua

3. Skenario Ketiga 60:40

Selanjutnya klasifikasi skenario ketiga 60:40 dengan data *testing* sebesar 40% dengan jumlah data *testing* sebanyak 384 data dan data *training* sebanyak 575 data menghasilkan nilai *accuracy* sebesar 0,8802 seperti yang terlihat pada Gambar 4.10.

```
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy:2f}')

Accuracy: 0.880208
```

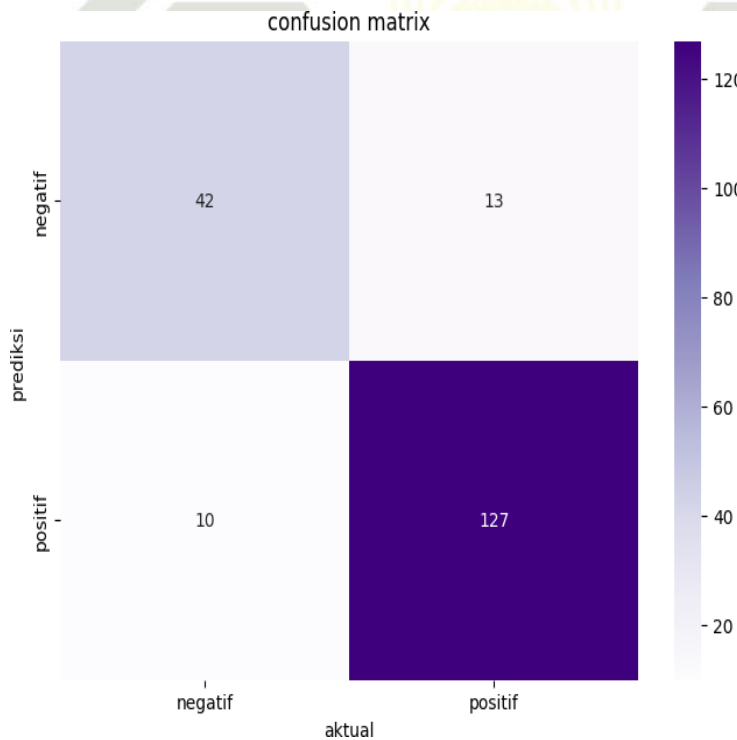
**Gambar 4.10.** Klasifikasi Skenario Ketiga

#### 4.6 Hasil Confusion Matrix

*Confusion matrix* ialah tahap testing untuk menentukan kelas asli dan prediksi dari hasil klasifikasi yang telah dilakukan. Pengujian *confusion matrix* berguna untuk mengetahui bagaimana keakuratan prediksi dari algoritma SVM terhadap data yang telah dilakukan. Dengan *confusion matrix* kita dapat melakukan penjumlahan unuk mengetahui nilai dari *accuracy*, *precision*, *recall* dan *f1-score*.

1. Skenario Pertama 80:20

Berikut ini adalah hasil *confusion matrix* dari setiap skenario yang telah dilakukan. Pada skenario pertama perbandingan 80:20, hasil *confusion matrix* menunjukkan bahwa model berhasil memprediksi 137 positif dan 55 negatif. dapat dilihat pada Gambar 4.11.



**Gambar 4.11.** Confusion matrix Skenario Pertama

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{127 + 42}{192} = \frac{169}{192} = 0,8802 \quad (4.1)$$



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.

$$Precision = \frac{TP}{TP + FP} = \frac{127}{137} = 0,93 \quad (4.2)$$

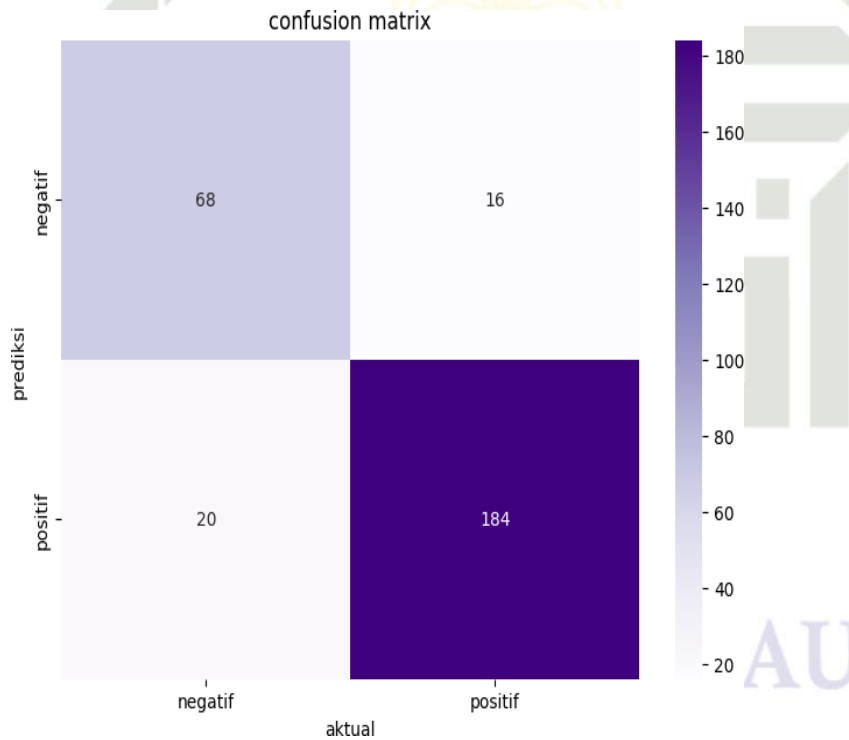
$$Recall = \frac{TP}{TP + FN} = \frac{127}{140} = 0,91 \quad (4.3)$$

$$F1\ Score = 2 \frac{Recall \times Precision}{Recall + Precision} = 2 \frac{0,837}{1,83} = 0,92 \quad (4.4)$$

Selanjutnya dilakukan perhitungan dengan menggunakan rumus dari *confusion matrix* pada skenario pertama menghasilkan sebagai berikut. *Accuracy*: 0,8802 *Precision*: 0,93 *Recall*: 0,91 *F1-score*: 0,92

2. Skenario Kedua 70:30

Selanjutnya pada skenario kedua pada perbandingan 70:30 hasil *confusion matrix* menunjukkan bahwa model berhasil memprediksi 204 positif dan 84 negatif seperti yang terlihat pada Gambar 4.12.



Gambar 4.12. Confusion matrix Skenario Kedua

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{184 + 68}{288} = \frac{252}{288} = 0,875 \quad (4.5)$$

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

$$Precision = \frac{TP}{TP + FP} = \frac{184}{204} = 0,90 \quad (4.6)$$

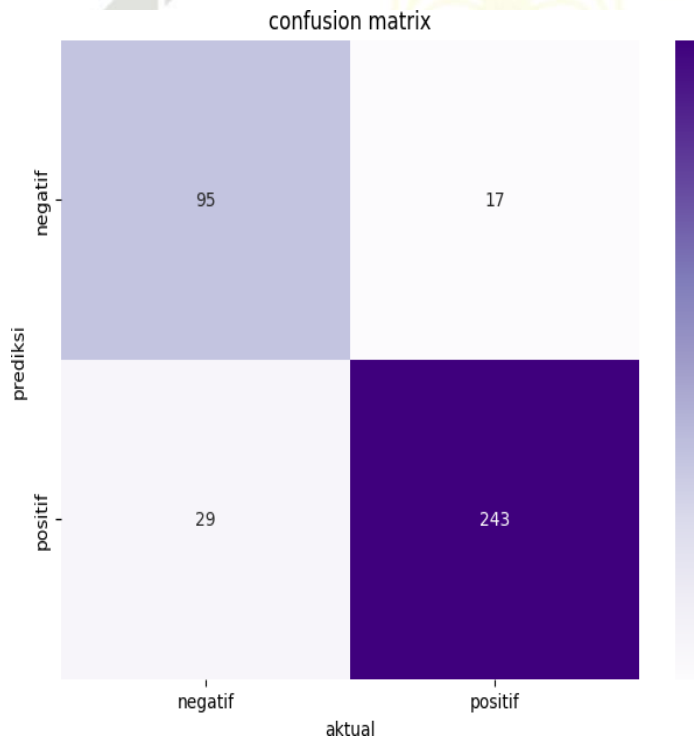
$$Recall = \frac{TP}{TP + FN} = \frac{184}{200} = 0,92 \quad (4.7)$$

$$F1\ Score = 2 \frac{Recall \times Precision}{Recall + Precision} = 2 \frac{0,828}{1,82} = 0,91 \quad (4.8)$$

Selanjutnya dilakukan perhitungan dengan menggunakan rumus dari *confusion matrix* pada skenario kedua menghasilkan sebagai berikut. *Accuracy*: 0,875 *Precision*: 0,90 *Recall*: 0,92 *F1-score*: 0,91

3. Skenario Ketiga 60:40

Selanjutnya pada skenario ketiga pada perbandingan 60:40 hasil *confusion matrix* menunjukkan bahwa model berhasil memprediksi 272 positif dan 112 negatif seperti yang terlihat pada Gambar 4.13.



Gambar 4.13. *Confusion matrix* Skenario Ketiga

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{338}{384} = 0,88 \quad (4.9)$$



**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.

$$Precision = \frac{TP}{TP + FP} = \frac{243}{272} = 0,89 \quad (4.10)$$

$$Recall = \frac{TP}{TP + FN} = \frac{243}{260} = 0,93 \quad (4.11)$$

$$F1\ Score = 2 \frac{Recall \times Precision}{Recall + Precision} = 2 \frac{0,82}{1,82} = 0,90 \quad (4.12)$$

Selanjutnya dilakukan perhitungan dengan menggunakan rumus dari *confusion matrix* pada skenario ketiga menghasilkan sebagai berikut. *Accuracy:0.88 Precision:0,89 Recall:0,93 F1-score:0,91*

**4.7 Hasil**

Berdasarkan hasil pengujian yang dilakukan terhadap 3 skenario, maka didapatkan hasil percobaan dengan skenario pertama 80:20 memiliki nilai hasil pengujian paling baik yaitu dengan nilai *accuracy* 88%, *precision* 93% dan *recall* 91%. Dari total 192 data *testing* yang digunakan, data aktual yang bernilai positif sebanyak 140 data sedangkan data aktual yang bernilai negatif sebanyak 52 data. Model memprediksi data yang bernilai positif sebanyak 127 data dan memprediksi data bernilai negatif sebanyak 42 data.