

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1 Deskripsi Penelitian**

Pada bagian ini melakukan evaluasi data berdasarkan pembagian *data set*. Jumlah data yang digunakan sebanyak 2.302 data komentar. Untuk membandingkan akurasi dalam mengklasifikasikan data, peneliti melakukan beberapa pembagian data *training* dan data *testing* secara random diantaranya.

#### **4.2 Evaluasi Hasil Penelitian**

##### **4.2.1 Pengumpulan Data**

Pada tahap ini, langkah yang dilakukan adalah pengumpulan data dengan proses *crawling* menggunakan Google Colab *Python*. Data yang diperlukan untuk penelitian ini berasal dari komentar-komentar YouTube. *Data set* tersebut diperoleh dari video berjudul "Klarifikasi (part 1)" yang diunggah oleh saluran YouTube Mgdalenaf pada tanggal 5 April 2023. Proses pengumpulan data melibatkan kunjungan ke video YouTube yang telah menerima banyak komentar terkait mengarah ke cyberbullying yang di unggah pada saluran YouTube.

Pengumpulan data dari komentar-komentar di video YouTube ini akan diambil dalam rentang waktu antara 05 April 2023 hingga 11 November 2023. Melalui proses *crawling* pada komentar video YouTube, terkumpul sebanyak 2.351 data komentar dari total 2.900 komentar proses *crawling* dilakukan pada tanggal 13 November 2023. Hasil data yang diperoleh dari proses *crawling* ini akan disimpan dalam format berkas *.csv*. Berikut data hasil *Crawling* yang di dapatkan.

publishedAt	authorDisplayName	textDisplay	likeCount
2023-11-17T23:43:59Z	Mu'a Brothers	CHANNEL FOOD VLOGGER MENDING HOBBY MAKAN SIH, DATENG-REVIEW-BORONG-BAGI GRATIS 🍕	1
2023-11-07T09:20:19Z	cha	Buset, gada yang nanya gimna konsep per media an lu magdaa	0
2023-11-04T23:21:04Z	withacaveat	!quoteKejijesama review sukarela!quote, bukan kerja sama Mbak e, anda butuh konten, anda minta izin, diberi izin bila pelaku usaha berker	0
2023-11-03T18:17:49Z	hahahhhi	taiii asuuuu hangsatittt magdangentodddd	1
2023-10-31T10:03:43Z	Ragit DP	PEMBELAAN DIRI BUKAN KELARIFIKASI 🍕	1
2023-10-30T15:43:57Z	Murdani Yusuf	Gk nanya SOP	0
2023-10-28T03:18:09Z	Arief Kurniawan	Pak Bondan yang jelas pioneer reviewer makanan aja ga sombong. Lah lu anak baru kemaren udah berasa food vlogger kawakan	0
2023-10-26T06:32:11Z	Yudis Arya	si ngemis makan 🍕	0
2023-10-25T10:13:47Z	@YUTTA100	love you magda 🍕	0
2023-10-24T17:01:16Z	kuyang sumatra	Play fucking victim	0
2023-10-22T20:05:52Z	CaptNemo	mh cewe gak ada bedanya sama ormas PP kocak 🍕	1
2023-10-20T22:48:50Z	Uba Ailin	Kalo gak minta di review lu mau apa	0
2023-10-20T12:42:02Z	Risman Rismana	Si paling ngartis, si paling populer	0
2023-10-16T18:03:35Z	Dan+ Xisaaar	Rampok makanan	0
2023-10-14T15:06:09Z	Kadek Gunawan Bali	Review sukarela sama dengan makan gratis 🍕🍕🍕🍕	0
2023-10-14T05:53:45Z	Arief Fakhruddin	Klo org ga mau direview ama loa sdh dong ga ush dibawa2 ke media, kan ada hak jg dong untuk notak. 🍕	0
2023-10-12T16:39:27Z	Haru Kurnia Putra	Unsubscribe rame2	0
2023-10-10T03:41:01Z	08 Wahyu Satyawan	Alasan pasti	1
2023-10-10T03:09:27Z	GhouiGuy	Makan gratis 🍕	1

Gambar 4. 1 Data Hasil Crawling dengan Google Colab

#### 4.2.2 Preprocessing Text

Peneliti melakukan serangkaian langkah *preprocessing text* pada kolom "textDisplay" menggunakan beberapa tahapan proses berikut hasilnya :

##### a. Cleaning

Tabel 4. 1 Cleaning

DATA ASLI	CLEANING
Review sukarela sama dengan makan gratis 🍕 🍕 🍕 🍕	Review sukarela sama dengan makan gratis
SALAH LANGKAH KLO MELAKUKAN PEMBELAAN RENDAHAN HATI seperti ini. Makin TERLIHAT NGOTOT PEMBENERAN DIRI.	SALAH LANGKAH KLO MELAKUKAN PEMBELAAN RENDAHAN HATI seperti iniMakin TERLIHAT NGOTOT PEMBENERAN DIRI

## b. Case Folding

Tabel 4. 2 Case Folding

CLEANING	CASE FOLDING
Review sukarela sama dengan makan gratis	review sukarela sama dengan makan gratis
SALAH LANGKAH KLO MELAKUKAN PEMBELAAN RENDAHAN HATI seperti ini Makin TERLIHAT NGOTOT PEMBENERAN DIRI	salah langkah klo melakukan pembelaan rendahan hati seperti inimakin terlihat ngotot pembeneran diri

## c. Tokenization

Tabel 4. 3 Tokenization

CASE FOLDING	TOKENIZATION
review sukarela sama dengan makan gratis	['review', 'sukarela', 'sama', 'dengan', 'makan', 'gratis']
salah langkah klo melakukan pembelaan rendahan hati seperti inimakin terlihat ngotot pembeneran diri	['salah', 'langkah', 'klo', 'melakukan', 'pembelaan', 'rendahan', 'hati', 'seperti', 'inimakin', 'terlihat', 'ngotot', 'pembeneran', 'diri']

## d. Filtering atau Stop Removal

Tabel 4. 4 Filtering atau Stop removal

TOKENIZATION	STOP REMOVAL
['review', 'sukarela', 'sama', 'dengan', 'makan', 'gratis']	['review', 'sukarela', 'makan', 'gratis']

['salah', 'langkah', 'klo', 'melakukan', 'pembelaan', 'rendahan', 'hati', 'seperti', 'inimakin', 'terlihat', 'ngotot', 'pembeneran', 'diri']	['salah', 'langkah', 'klo', 'pembelaan', 'rendahan', 'hati', 'inimakin', 'ngotot', 'pembeneran']
--	---

#### e. Steaming Data

Tabel 4. 5 Steaming

STOP REMOVAL	STEAMING DATA
['review', 'sukarela', 'makan', 'gratis']	review sukarela makan gratis
['salah', 'langkah', 'klo', 'pembelaan', 'rendahan', 'hati', 'inimakin', 'ngotot', 'pembeneran']	salah langkah klo bela rendah hati inimakin ngotot pembeneran

#### 4.2.3 Pelabelan

Setelah melalui tahap *preprocessing*, langkah berikutnya adalah melakukan pelabelan pada data menjadi kategori *cyberbullying* dan *non\_cyberbullying* sebelum dilakukan proses klasifikasi. Oleh karena itu, data yang akan digunakan untuk melatih model perlu diberikan label terlebih dahulu. Dalam penelitian ini, digunakan dua metode pelabelan untuk mengklasifikasikan komentar YouTube.

Metode pertama pelabelan data melibatkan pemberian label secara manual dengan membaca dan mengevaluasi komentar satu per satu, sementara metode kedua yaitu menggunakan Vader Lexicon di *Python*.

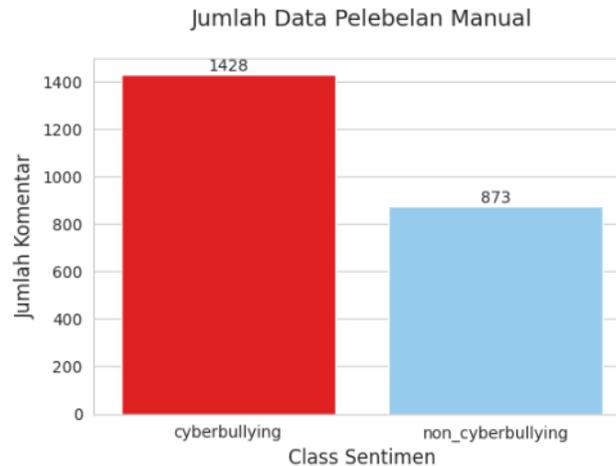
### a. Pelabelan Secara Manual

Data akan diberi label hanya dalam dua kelas sentimen: *cyberbullying* dan *non\_cyberbullying*. Pemberian label secara manual dengan membaca dan mengevaluasi ulasan satu per satu. Dalam penelitian ini, terlibat 3 partisipan yaitu, Deani Intan Sari, Putri Oktavia Legina dan Ridha Putri untuk melakukan penilaian subjektif terhadap konten komentar dan memberikan label sentimen *cyberbullying* atau *non\_cyberbullying* berdasarkan penilaian mereka. Partisipan melakukan pelabelan secara manual menggunakan google sheet. Berikut hasil data pelabelan data secara manual.

	A	B	C	D	E	F	G
1	textDisplay	cleasing	case_folding	tokenize	filtering/stopword	steaming_data	klasifikasi
2	CHANNEL FOO CHANNEL FOO	channel food vlo	[channel, 'food'	[channel, 'food'	channel food vlogger	mending hobby makan sih datengreviewborongbagi gratis	cyberbullying
3	Buset, gaada ya Buset gaada yar	buset gaada yar	[buset, 'gaada'	[buset, 'gaada'	buset gaada nanya gmna konsep media an lu magdaa		non_cyberbullying
4	&quot;Kerjasama r quot;Kerjasama r	quot;kerjasama r	[quot;kerjasama	[quot;kerjasama	review sukarelaquot kerja mbak e butuh konten izin izin usaha kenan kerjasama		non_cyberbullying
5	talii asuuuu ban talii asuuuu ban	talii asuuuu ban	[talii, 'asuuuu'	[talii, 'asuuuu'	'talii, 'asuuuu, 'talii asuuuu bangsat!!!! maqdgantododddd		cyberbullying
6	PEMBELAAN DI PEMBELAAN DI	pembelaan diri t	[pembelaan, 'di	[pembelaan, 'di	bela katarifikasi		non_cyberbullying
7	Gk nanya SOP Gk nanya SOP	gk nanya sop	[gk, 'nanya, 'so	[gk, 'nanya, 'so	gk nanya sop		non_cyberbullying
8	Pak Bondan yan Pak Bondan yan	pak bondan yan	[pak, 'bondan'	[bondan, 'jama	bondan pioneer reviewer makan aja ga sobong lu anak kemaren udah asa food vlogger kawak		cyberbullying
9	si ngemis makar si ngemis makar	si ngemis makar	[si, 'ngemis, 'mi	[si, 'ngemis, 'mi	si ngemis makan		cyberbullying
10	love you magda love you magda	love you magda	[love, 'you, 'ma	[love, 'you, 'ma	love you magda		non_cyberbullying
11	Play fucking victi Play fucking victi	play fucking victi	[play, 'fucking, 'i	[play, 'fucking, 'i	play fucking victim		cyberbullying
12	nih cewe gak ad nih cewe gak ad	nih cewe gak ad	[nih, 'cewe, 'gal	[nih, 'cewe, 'gal	nih cewe gak beda ormas pp kocak		cyberbullying
13	Kalo gak minta c Kalo gak minta c	kalo gak minta d	[kalo, 'gak, 'min	[kalo, 'gak, 'rev	kalo gak reviuur lu		non_cyberbullying
14	Si paling ngartis. Si paling ngartis	si paling ngartis	[si, 'paling, 'nga	[si, 'ngartis, 'si,	si ngartis si populer		cyberbullying
15	Rampok makani Rampok makani	rampok makana	[rampok, 'maka	[rampok, 'maka	rampok makan		cyberbullying
16	Review sukarela Review sukarela	review sukarela	[review, 'sukare	[review, 'sukare	review sukarela makan gratis		non_cyberbullying
17	Klo org ga mau klo org ga mau	klo org ga mau c	[klo, 'org, 'ga	[klo, 'org, 'ga	'klo org ga dreview ama loe sdh ga ush bawia media hak jg nolak		non_cyberbullying
18	Unsubscribe ran Unsubscribe ran	unsubscribe ram	[unsubscribe, 'i,	[unsubscribe, 'i,	unsubscribe rame		cyberbullying
19	Alasan pasti Alasan pasti	alasan pasti	[alasan, 'pasti]	[alasan]	alasan		non_cyberbullying
20	Makan gratis Makan gratis	makan gratis	[makan, 'gratis]	[makan, 'gratis]	makan gratis		non_cyberbullying

Gambar 4. 2 Data Hasil Pelabelan Manual

Pada gambar di bawah diketahui bahwa hasil sentimen pada komentar YouTube lebih banyak komentar *cyberbullying* dari pada *non\_cyberbullying*. Untuk hasil sentiment dengan Pelabelan Manual didapatkan kelas *cyberbullying* sebanyak 1428 data komentar, sedangkan *non\_cyberbullying* sebesar 873 data komentar.



Gambar 4. 3 Jumlah Data Pelebelan Manual

#### b. Pelabelan Menggunakan Vader Lexicon

Metode kedua pelabelan untuk kelas sentimen dilakukan otomatis menggunakan *package* NLTK dalam bahasa pemrograman *Python*, dengan bantuan modul Vader Lexicon. Dalam proses pelabelan menggunakan Vader Lexicon ini, terdapat penggunaan sistem *compound* yang digunakan untuk menghitung nilai yang menggambarkan konotasi yang berkaitan dengan '*cyberbullying*' maupun '*non\_cyberbullying*'.

*Compound* adalah nilai yang mengukur secara keseluruhan, semua skor yang telah dinormalisasi ke dalam rentang antara -1 dan +1. Ketika nilai *compound* berada di bawah 0 ( $\text{compound} < 0$ ), dianggap sebagai sentimen *cyberbullying*, sementara nilai *compound* di atas 0 ( $\text{compound} \geq 0$ ) dianggap sebagai sentimen *non\_cyberbullying*.

Jika kata *non\_cyberbullying* lebih dominan, kalimat diklasifikasikan sebagai sentimen *non\_cyberbullying*, sebaliknya jika kata *cyberbullying* lebih dominan, maka kalimat diklasifikasikan sebagai sentimen *cyberbullying*.

Kode berikut untuk proses pelabelan data:

**CODE**

```
# install library, punkt tokenizer dan vader lexicon
!pip install nltk
nltk.download('punkt')
nltk.download('vader_lexicon')

#pelebelan
data = SentimentIntensityAnalyzer()
df['steaming_data'] = df['steaming_data'].fillna("")
labels = []
scores = []
for text in df['steaming_data']:
    sentiment_scores = data.polarity_scores(text)
    compound_score = sentiment_scores['compound']
    scores.append(compound_score)
    if compound_score > 0:
        label = 'non_cyberbullying'
    else:
        label = 'cyberbullying'
    labels.append(label)
df['sentimen_score'] = scores
df['sentimen'] = labels
```

Berikut hasil outputnya :

	steaming_data	sentimen_score	sentimen
0	channel food vlogger mending hobby makan sih d...	0.0516	non_cyberbullying
1	buset gaada nanya gmna konsep media an lu magdaa	0.0000	cyberbullying
2	quotkerjasama review sukarelaquot kerja mbak e...	0.0000	cyberbullying
3	taiiii asuuuu bangsatTTTT maqdangentodddds	0.0000	cyberbullying
4	bela klarifikasi	0.0000	cyberbullying
5	gk nanya sop	0.0000	cyberbullying
6	bondan pioneer reviewer makan aja ga sombong l..	0.0000	cyberbullying
7	si ngemis makan	0.0000	cyberbullying
8	love you magda	0.6369	non_cyberbullying
9	play fucking victim	0.0018	non_cyberbullying

**Gambar 4. 4 Data Hasil Pelabelan Vader Lexicon**

Dapat dilihat gambar di bawah diketahui bahwa hasil sentimen pada komentar YouTube lebih banyak komentar *cyberbullying* dari pada *non\_cyberbullying*. Untuk hasil sentimen dengan perhitungan Vander Lexicon didapatkan kelas *cyberbullying* sebesar 1868 data komentar, sedangkan *non\_cyberbullying* sebesar 434 data komentar.

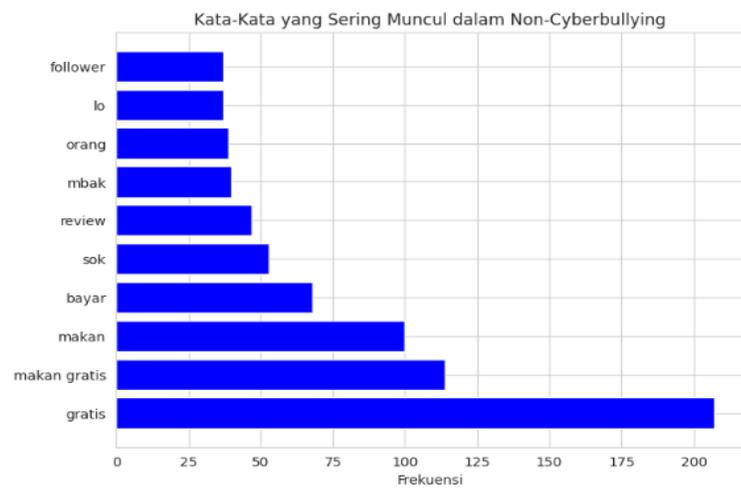
Dapat diperoleh informasi mengenai jumlah komentar yang terlabel sebagai *cyberbullying* dan *non-cyberbullying* menggunakan kode berikut ini.

```
sentimen_count = data['sentimen'].value_counts()
```

Berikut hasil outputnya :







**Gambar 4. 7 Wordcloud dan Diagram Sentimen Non-Cyberbullying**

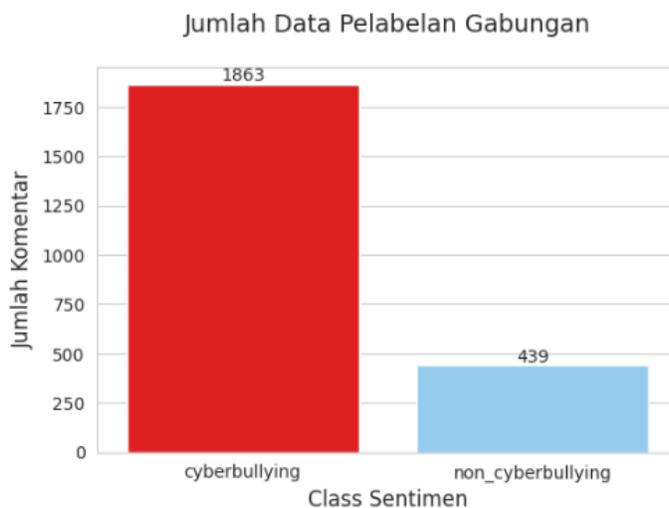
Pada Gambar 4.8 merupakan wordcloud dan diagram *sentimen non-cyberbullying* pada komentar akun youtube Mgdalenaf. Kata yang sering muncul meliputi “gratis”, “makan gratis”, “makan”, “bayar”, “sok”, “review”, “mbak”, “orang”, “lo”, “follower”. *Wordcloud* tersebut memberikan gambaran visual tentang kata-kata yang dominan dalam konteks sentimen *non-cyberbullying* pada komentar akun tersebut.

#### **4.2.5 Penggabungan Data Pelabelan Manual dan Vader Lexicon**

Pada penelitian ini, dilakukan penggabungan hasil pelabelan sentimen dari dua sumber berbeda, yaitu pelabelan manual untuk 461 data komentar dan pelabelan analisis sentimen otomatis untuk 1.841 data komentar. Gabungan tersebut membentuk satu set data yang terdiri dari 2.302 komentar secara keseluruhan.

Penting untuk memahami bahwa penggabungan hasil pelabelan sentimen dilakukan untuk memanfaatkan kekuatan kedua metode tersebut. Pelabelan manual memberikan wawasan langsung dari pandangan manusia, sementara analisis sentimen otomatis menggunakan pendekatan Vader Lexicon memberikan efisiensi dalam menangani jumlah data yang

lebih besar. Berikut visualisasi sentimen penggabungan data pelabelan manual dan vader lexicon :



**Gambar 4. 8 Jumlah Data Pelabelan Gabungan**

Pada gambar di atas, diketahui bahwa hasil sentimen dari penggabungan data pelabelan manual dan Vader Lexicon menunjukkan adanya sentimen *cyberbullying* sebanyak 1.863 sentimen, sedangkan sentimen *non-cyberbullying* terdapat sebanyak 439 sentimen.

#### **4.2.6 Pembagian Data Training dan Data Testing**

Sebelum melanjutkan ke tahap klasifikasi, langkah awalnya adalah membagi data komentar menjadi dua bagian: data latih dan data uji. Data latih digunakan untuk mengembangkan model klasifikasi, di mana model ini akan mempelajari pola-pola dari data tersebut. Pengetahuan yang diperoleh oleh model dari data latih akan digunakan untuk memprediksi kelas sentimen dari data baru yang belum pernah dilihat sebelumnya. Semakin besar jumlah data yang digunakan untuk melatih model, maka algoritma akan semakin baik dalam memahami pola-pola dalam data tersebut.

Data uji dimanfaatkan untuk mengevaluasi seberapa baik algoritma klasifikasi dapat melakukan prediksi yang tepat. Dalam penelitian ini, data dibagi menjadi 80% untuk data latih dan 20% untuk data uji. Pemisahan data menjadi data *training* dan data *testing* dilakukan dengan menggunakan kode berikut, dan berikut adalah hasilnya:

CODE
<pre>#membagi data menjadi data training dan testing dengan test_size = 0.20 dan random state nya 0 from sklearn.model_selection import train_test_split X_train, X_test, y_train, y_test = train_test_split(data['steaming_data'], data['klasifikasi'], test_size = 0.20, random_state = 0)</pre>

```
Data training (X_train):
390      tau kalo review food nya ga jujur tp pake baya...
1570     dri editing video nya nya ting pikir celah dri...
254                                     chinaaaaa mintak maaf
575                                           iya ka klarifikasi gratis gabayar
2191                                         semangatt kakk
933                                           ni kolak sisa buka gagratis
721     tutorial dimaafin dimaklumin orangkesalahan ba...
1968     tutorial klarifikasi maaf cuman channel tretan...
303      udah gilak klarifikasi aj part part nya hhhh t...
961      waduhhh dlu udh bnyak bgt tmn umkm yg curhat k...
Name: steaming_data, dtype: object
```

```
Data testing (X_test):
124      pungut biaya apa sekalijadi makan nya bikin em...
1651     salah tafsir mbakgini lo konsepnyakalo rumah m...
1427                                           semangat ka
1961                                           klarifikasi maaf knp pembelaan
983      raffi ahmad gigi jajarans makan bayar bantu pr...
1577                                           kocak
486                                           pas sdh bully haq
1330                                           orang kalo salah ya bela
1727                                           nonton chanel belah
1969                                           bisnis tetep bisnis yh ka bskquot bayar
Name: steaming_data, dtype: object
```

**Gambar 4. 9 Hasil Pembagian Data Training dan Testing**

#### 4.2.7 Pembobotan TF-IDF

Gambar yang ditampilkan di bawah adalah representasi TF-IDF dari dokumen-dokumen pada dataset. Setiap baris mewakili sebuah dokumen, dan setiap kolom mewakili kata dalam kamus (fitur) yang ditemukan dalam dataset.

Masing-masing nilai dalam *matriks* menunjukkan bobot TF-IDF untuk kata tertentu dalam dokumen tertentu. Bobot ini menggambarkan seberapa pentingnya kata tersebut dalam dokumen tersebut relatif terhadap keseluruhan dataset.

(0, 3747)	0.4888446627068671
(0, 299)	0.2128063676461445
(0, 2610)	0.2877663542607402
(0, 3524)	0.32861615247240844
(0, 1523)	0.3948872206771763
(0, 1054)	0.2195555557811426
(0, 2520)	0.2335998971488615
(0, 1034)	0.2607935592119646
(0, 2958)	0.2599875030113636
(0, 1575)	0.23412706052043364
(0, 3407)	0.2753444416163856
(1, 2025)	0.11289765919658495
(1, 3345)	0.23970812149442755
(1, 3554)	0.18624852779882362
(1, 1286)	0.1576372385162313
(1, 568)	0.30177278776597055
(1, 2727)	0.23603268023873486
(1, 3489)	0.30177278776597055
(1, 3660)	0.18529630792879703
(1, 902)	0.30177278776597055
(1, 868)	0.5500859818363372
(1, 299)	0.13136927886934038
(1, 2520)	0.4326165194938048
(2, 1983)	0.3067213708290902
(2, 2199)	0.6548925928152926
:	:
(1838, 1618)	0.1160310881820605
(1838, 3682)	0.11319316506371295
(1838, 1215)	0.10399973338066316
(1838, 2647)	0.1252245198651103
(1838, 2952)	0.10490680603222775
(1838, 97)	0.13499200227019176
(1838, 866)	0.1145677551525633
(1838, 1513)	0.14094079513355215

Gambar 4. 10 Hasil Tfidf Training

Dalam representasi sparse matrik, seperti yang ditampilkan setiap baris mewakili sebuah dokumen dan setiap kolom mewakili suatu kata atau

fitur. Contoh pada hasil (0, 3747) Nilai TF-IDF untuk kata dengan indeks 3747 pada dokumen 0 adalah 0.4888446627068671. Dengan kata lain, pada dokumen pertama (indeks 0), kata yang memiliki indeks 3747 dalam kamus memiliki nilai TF-IDF sebesar 0.4888446627068671. Nilai ini menunjukkan seberapa penting kata tersebut dalam dokumen tersebut, relatif terhadap keseluruhan dataset.

Apabila nilai ini tinggi, kata tersebut dianggap penting dalam dokumen tersebut. Sebaliknya, nilai rendah menunjukkan bahwa kata tersebut mungkin tidak terlalu penting atau mungkin sering muncul di seluruh dataset (Wati, Ernawati, and Rachmi 2023).

### 4.3 Analisis Hasil

Dalam penggunaan metode *Naive Bayes* untuk deteksi *cyberbullying*, terdapat beberapa skenario yang dapat dipertimbangkan dengan menggunakan pelabelan manual, otomatis (Vader Lexicon), dan penggabungan keduanya. Proses ini bertujuan untuk mengevaluasi penggunaan metode *Naive Bayes*. Berikut tahapan yang akan dibahas :

1. Pelabelan Data Dengan Manual
2. Pelabelan Data Dengan Otomatis
3. Pelabelan Data Dengan Penggabungan

#### 4.3.1 Skenario Pertama Pelabelan Data Manual

Pada skenario pertama, digunakan data hasil *preprocessing* yang terdiri dari 2.302 data komentar. Proses pembagian data *training* dan data *testing* dilakukan dengan proporsi 80% untuk data *training* (dalam hal ini, sekitar 1.841 data) dan 20% untuk data *testing* (sekitar 461 data). Penting untuk dicatat bahwa pembagian data *testing* dan *training* dilakukan secara acak untuk memastikan representativitas dalam mengukur kinerja model *Naive Bayes*. Dengan penggunaan dataset yang telah diolah sebanyak 2.302 komentar, skenario ini memungkinkan untuk melatih dan menguji model

dalam konteks pelabelan sentimen manual. Berikut hasil analisisnya dapat di lihat pada gambar di bawah ini :

```

Akurasi Model Naive Bayes : 0.6594360086767896
confusion_matrix:
[[263  11]
 [146  41]]
=====

Classification Report:
              precision    recall  f1-score   support

     0.0       0.64      0.96      0.77       274
     1.0       0.79      0.22      0.34       187

 accuracy              0.66       461
 macro avg              0.72       461
 weighted avg           0.70       461

Akurasi Cross-Validation: 67.03%

```

**Gambar 4. 11 Hasil Pelabelan Data Manual**

Berdasarkan hasil evaluasi model dengan *Confusion Matrix*, dapat dilihat bahwa True Positives (TP): 41, True Negatives (TN): 263, False Positives (FP): 11, False Negatives (FN): 146

Dengan demikian, model berhasil mengidentifikasi 41 sentimen positif secara tepat (TP) dan 263 sentimen negatif dengan benar (TN). Namun, terdapat 11 kasus di mana sentimen seharusnya negatif, tetapi model salah mengklasifikasikannya sebagai positif (FP). Selain itu, terdapat 146 kasus di mana sentimen seharusnya positif, tetapi model salah mengklasifikasikannya sebagai negatif (FN).

Dalam *Classification Report*, Akurasi (*Accuracy*): Seberapa baik model dapat mengklasifikasikan secara benar seluruh sentimen (*non-cyberbullying* dan *cyberbullying*). Dalam hasil akurasi sekitar 66%.

*Precision*, seberapa tepat model dalam mengklasifikasikan sentimen *non-cyberbullying* dan *cyberbullying*. Dalam kasus ini *precision* untuk sentimen *non-cyberbullying* adalah sekitar 0.79, dan untuk sentimen *cyberbullying* adalah sekitar 0.64.

*Recall*, sejauh mana model dapat mengenali seluruh sentimen *non-cyberbullying* dan *cyberbullying* yang seharusnya diidentifikasi. Dalam kasus ini, *recall* untuk sentimen *non-cyberbullying* adalah sekitar 0.22, dan untuk sentimen *cyberbullying* adalah sekitar 0.96.

*F1-score*, sebuah metrik yang mencerminkan keseimbangan antara *precision* dan *recall*. Dalam kasus Anda, *F1-score* untuk sentimen *non-cyberbullying* adalah sekitar 0.34, dan untuk sentimen *cyberbullying* adalah sekitar 0.77.

Dalam Akurasi *Cross-Validation*, model menunjukkan kemampuan yang baik untuk menggeneralisasi pola pada data baru yang belum pernah dilihat sebelumnya, dengan tingkat akurasi mencapai 67.03%. Ini memberikan indikasi positif bahwa model dapat memberikan performa yang memadai pada situasi dunia nyata dengan data baru.

#### **4.3.2 Skenario Kedua Pelabelan Data Otomatis**

Dalam Skenario kedua pelabelan otomatis (Vader Lexicon) dengan *Naive Bayes*, tetap mempertahankan jumlah total data komentar sebanyak 2.302. Proses pembagian data *training* dan *testing* tetap mengikuti proporsi 80% untuk data *training* (sekitar 1.841 data) dan 20% untuk data *testing* (sekitar 461 data), sesuai dengan skenario sebelumnya. Adapun proses pengambilan data secara acak dilakukan untuk memastikan representativitas. Berikut hasil analisisnya dapat di lihat pada gambar di bawah ini :

```

Akurasi Model Naive Bayes : 0.8590021691973969
confusion_matrix:
[[386  0]
 [ 65 10]]
=====

Classification Report:
              precision    recall  f1-score   support

 cyberbullying         0.86      1.00      0.92       386
non_cyberbullying      1.00      0.13      0.24        75

   accuracy                   0.86       461
  macro avg         0.93      0.57      0.58       461
 weighted avg         0.88      0.86      0.81       461

Akurasi Cross-Validation: 82.84%

```

**Gambar 4. 12 Hasil Pelabelan Data Otomatis**

Dengan akurasi model *Naive Bayes* sebesar 86%, terdapat hasil yang signifikan dalam mengenali sentimen *cyberbullying*. *Confusion matrix* menggambarkan distribusi hasil klasifikasi model, di mana terdapat 386 True Negatif (TN), 10 True Positif (TP), 0 False Positif (FP), dan 65 False Negatif (FN).

Dalam *Classification Report*, *precision* (presisi) untuk sentimen *cyberbullying* adalah 0.86, menunjukkan bahwa dari seluruh data yang diprediksi sebagai *cyberbullying*, 86% di antaranya adalah prediksi yang benar. *Recall* untuk sentimen *cyberbullying* adalah 1.00, yang menunjukkan bahwa model berhasil mengidentifikasi seluruh sentimen *cyberbullying* yang seharusnya diidentifikasi. Namun, untuk sentimen *non-cyberbullying*, *precision* adalah 1.00, tetapi *recall* yang rendah (0.13) menunjukkan bahwa model kurang mampu mengenali sebagian besar sentimen *non-cyberbullying* yang seharusnya diidentifikasi. *F1-score* untuk sentimen *cyberbullying* adalah 0.92, sedangkan untuk sentimen *non-cyberbullying* adalah 0.24.

Hasil validasi silang menunjukkan kinerja yang konsisten pada data yang belum dilihat sebelumnya, dengan akurasi mencapai 82.84%.

Secara keseluruhan, model *Naive Bayes* ini menunjukkan performa yang baik, terutama dalam mengidentifikasi sentimen *cyberbullying*. Namun, ada ruang untuk perbaikan dalam mengenali sentimen *non-cyberbullying* untuk meningkatkan keseimbangan antara *precision* dan *recall* untuk kedua kategori sentimen.

#### **4.3.3 Skenario Ketiga Pelabelan Data Penggabungan**

Dalam Skenario ketiga penggabungan pelabelan manual dan otomatis, dengan mengambil pendekatan untuk melibatkan informasi dari kedua sumber pelabelan, yaitu hasil pelabelan manual dan otomatis, dalam melatih dan menguji model menggunakan metode *Naive Bayes*. Jumlah total data komentar yang diolah dalam skenario ini tetap sebanyak 2.302.

Pembagian data *training* dan *testing* dipertahankan dengan proporsi 80% untuk data *training* (sekitar 1.841 data) dan 20% untuk data *testing* (sekitar 461 data), sesuai dengan konsistensi pada skenario sebelumnya. Proporsi ini dijaga agar memungkinkan perbandingan yang sebanding dalam evaluasi kinerja model. Berikut hasil analisisnya dapat di lihat pada gambar di bawah ini :

```

Akurasi Model Naive Bayes : 0.8026030368763557
confusion_matrix:
[[368  0]
 [ 91  2]]
=====

Classification Report:
              precision    recall  f1-score   support

 cyberbullying         0.80      1.00      0.89       368
non_cyberbullying      1.00      0.02      0.04        93

   accuracy                   0.80       461
  macro avg         0.90      0.51      0.47       461
 weighted avg         0.84      0.80      0.72       461

Akurasi Cross-Validation: 81.42%

```

**Gambar 4. 13 Hasil Pelabelan Data Penggabungan**

Hasil evaluasi model *Naive Bayes* menunjukkan akurasi sebesar 80.26%, yang menandakan bahwa model mampu mengklasifikasikan data dengan tepat sekitar 80.26% dari total data. *Confusion matrix* menggambarkan distribusi hasil klasifikasi model, di mana terdapat 368 True Negatif (TN), 2 True Positif (TP), 0 False Positif (FP), dan 91 False Negatif (FN).

Dalam *Classification Report*, *precision* (presisi) untuk sentimen *cyberbullying* adalah 0.80, menunjukkan bahwa dari seluruh data yang diprediksi sebagai *cyberbullying*, 80% di antaranya adalah prediksi yang benar. *Recall* untuk sentimen *cyberbullying* adalah 1.00, yang menunjukkan bahwa model berhasil mengidentifikasi seluruh sentimen *cyberbullying* yang seharusnya diidentifikasi. Namun, untuk sentimen *non-cyberbullying*, *precision* adalah 1.00, tetapi *recall* yang rendah (0.02) menunjukkan bahwa model kurang mampu mengenali sebagian besar sentimen *non-cyberbullying* yang seharusnya diidentifikasi. *F1-score* untuk sentimen *cyberbullying* adalah 0.89, sedangkan untuk sentimen *non-cyberbullying* adalah 0.04.

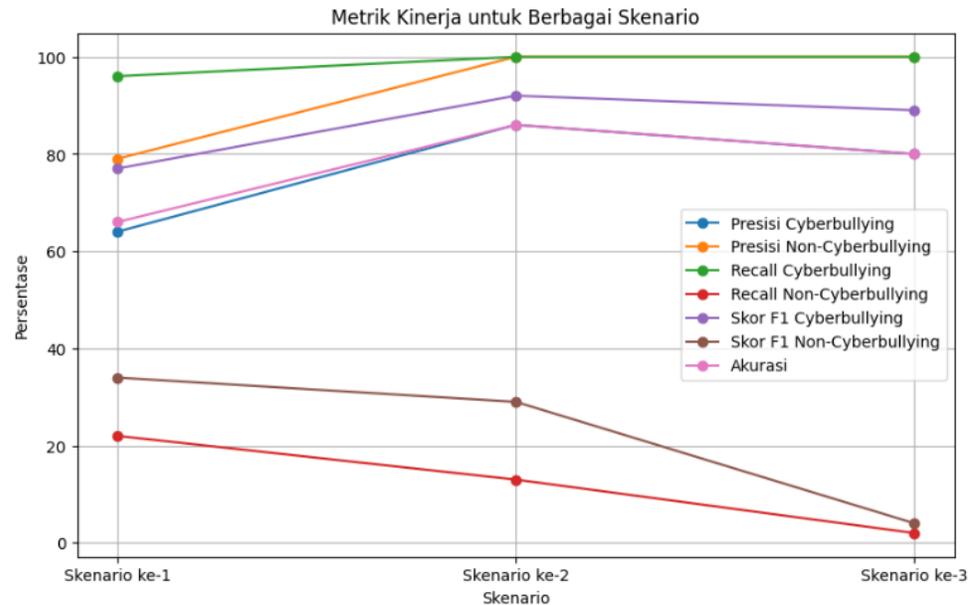
Akurasi *Cross-Validation*, sebesar 81.42% memberikan gambaran bahwa model memiliki kemampuan yang baik untuk menggeneralisasi pola pada data baru yang belum pernah dilihat sebelumnya. Penggabungan pelabelan manual dan otomatis diharapkan dapat memberikan perspektif yang lebih komprehensif terhadap sentimen dalam dataset.

#### 4.3.4 Hasil Klasifikasi

Berdasarkan evaluasi model *Naive Bayes* menggunakan data pelabelan manual, otomatis (*Vader Lexicon*), dan data penggabungan keduanya, dapat disimpulkan bahwa ketiga skenario menunjukkan perbandingan performa menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* dapat dilihat pada table perbandingan berikut :

Tabel 4. 6 Tabel Kinerja Skenario

Nilai	Skenario ke-1 (Data Pelabelan Manual)			Skenario ke-2 (Data Pelabelan Otomatis)			Skenario ke-3 (Data Pelabelan Penggabungan)		
	Cyber bullying	Non-Cyber bullying	Akurasi	Cyber bullying	Non-Cyber bullying	Akurasi	Cyber bullying	Non-Cyber bullying	Akurasi
Precision	64%	79%	66%	86%	100%	86%	80%	100%	80%
Recall	96%	22%		100%	13%		100%	2%	
F1-Score	77%	34%		92%	29%		89%	4%	



Gambar 4. 14 Diagram Nilai Accuracy

Dari hasil evaluasi model *Naive Bayes* pada tiga skenario, terlihat bahwa: Skenario ke-2 (Data Pelabelan Otomatis) memiliki kinerja yang sangat baik dalam mengidentifikasi sentimen *cyberbullying* dengan akurasi tertinggi (86%). Karena skenario ini memiliki *recall* dan *precision* yang tinggi untuk kedua kategori sentimen, menunjukkan bahwa model mampu dengan baik mengenali dan mengklasifikasikan komentar sebagai *cyberbullying* atau *non-cyberbullying*. Secara keseluruhan, skenario ini memberikan hasil yang optimal dalam menghadapi tantangan klasifikasi sentimen pada dataset.

Skenario ke-3 (Data Pelabelan Penggabungan) juga memberikan hasil yang baik, tetapi sedikit lebih rendah dibandingkan dengan Skenario ke-2. Meskipun akurasi totalnya 80%, model ini menunjukkan keseimbangan yang lebih baik antara kedua kategori sentimen, dengan *recall* dan *precision* yang baik untuk *cyberbullying* dan *non-cyberbullying*.

Skenario ke-1 (Data Pelabelan Manual) memiliki akurasi yang rendah (66%) dibandingkan dengan dua skenario lainnya dan kesulitan dalam mengenali sentimen *non-cyberbullying* (*recall* rendah). Hal ini dapat disebabkan oleh beberapa faktor, seperti subjektivitas dalam pelabelan manual atau kerumitan variasi dalam sentimen yang sulit diidentifikasi secara konsisten.