

## BAB 4

### HASIL PENELITIAN

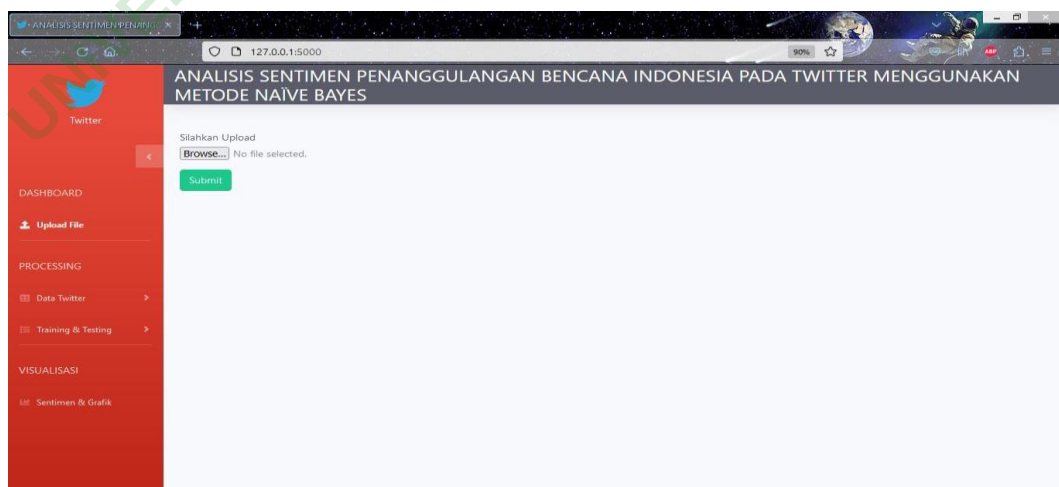
#### 4.1 RINGKASAN HASIL PENELITIAN

Hasil dari metode Naïve Bayes Classification dilakukan untuk menemukan hasil data dari Twitter untuk sentimen positif dan negatif yang diperoleh dari tanggal 1 juni 2020 –17 agustus 2022 dengan 2 kata kunci yaitu “Penanggulangan Bencana” dan “Bantuan Bencana” dengan jumlah keseluruhan 32.000 data baik *tweet* dan *retweet*. Kemudian dilanjutkan dengan tahap training data dengan jumlah 400 *tweet* data terdiri dari data *tweet* negatif 200 dan data *tweet* positif 200 yang telah dilabeli dengan manual, Kemudian tahap testing menggunakan 200 data yang telah dilabeli. Dengan adanya hasil penelitian ini harapannya memiliki dampak yang baik mengenai *tweet* penanggulangan bencana di indonesia yang bisa dijadikan bahan untuk evaluasi dari kinerja Badan Penanggulangan Bencana Indonesia agar lebih baik kedepannya.

#### 4.2 IMPLEMENTASI SISTEM

##### 4.2.1 Fitur Upload File

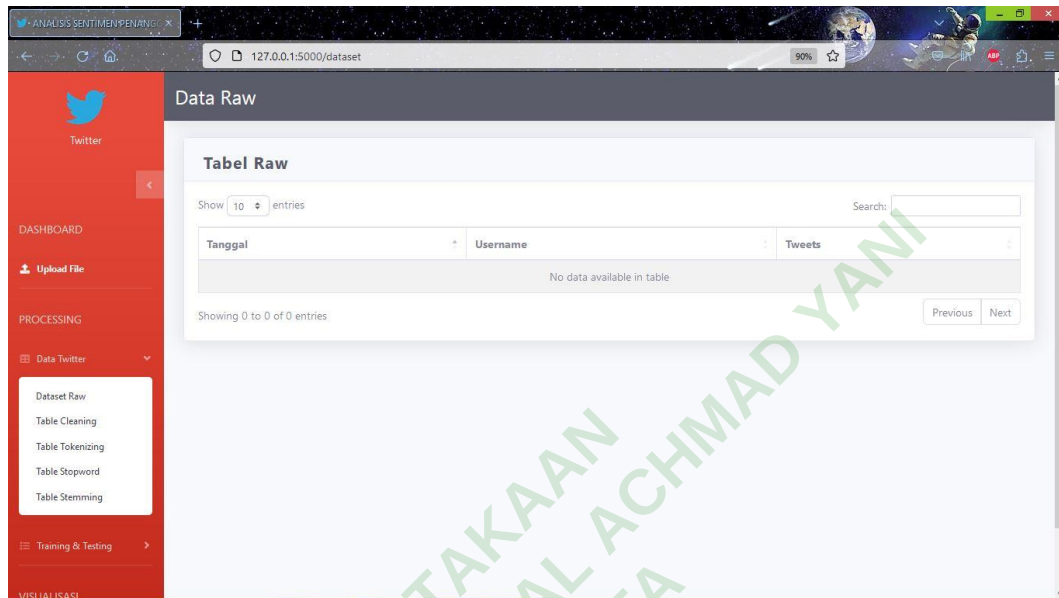
Pada fitur upload file terdapat tempat untuk mengirimkan file csv hasil crawling data ke sistem. Tampilan fitur tersebut sebagai berikut.



Gambar 4.1 Fitur Upload File

### 4.2.2 Fitur Dataset Raw

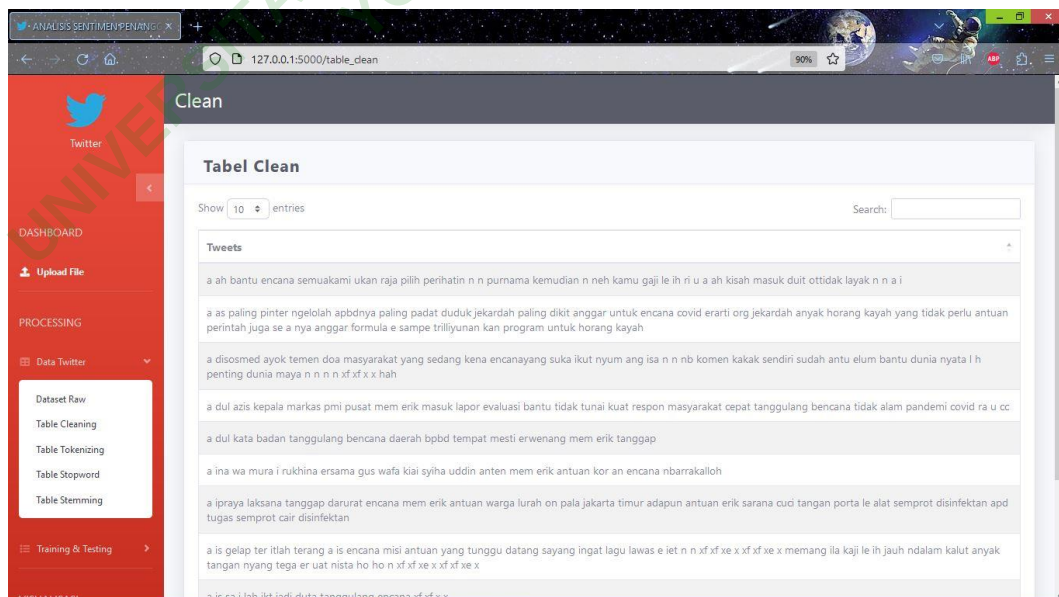
Pada fitur ini data yang telah di upload pada fitur upload file akan diproses untuk ditampilkan disini. Tampilan fitur tersebut dapat dilihat sebagai berikut.



Gambar 4.2 Fitur Dataset Raw

### 4.2.3 Fitur Table Cleaning

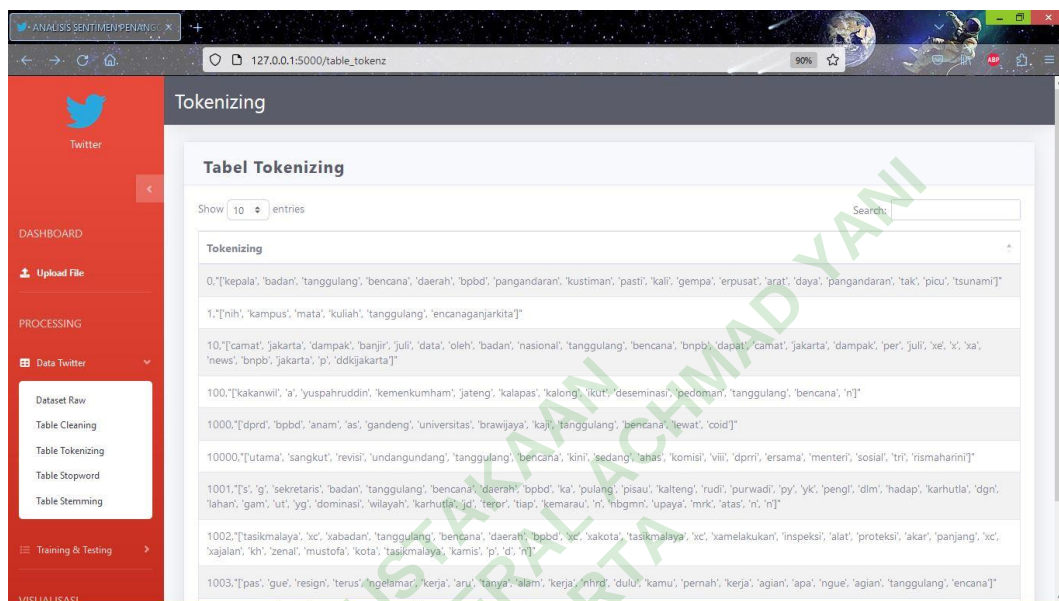
Pada fitur ini dataset yang kotor telah dilakukan pembersihan yang kemudian akan digunakan ke tahap tokenizing. Tampilannya sebagai berikut.



Gambar 4.3 Fitur Table Cleaning

#### 4.2.4 Fitur Table Tokenizing

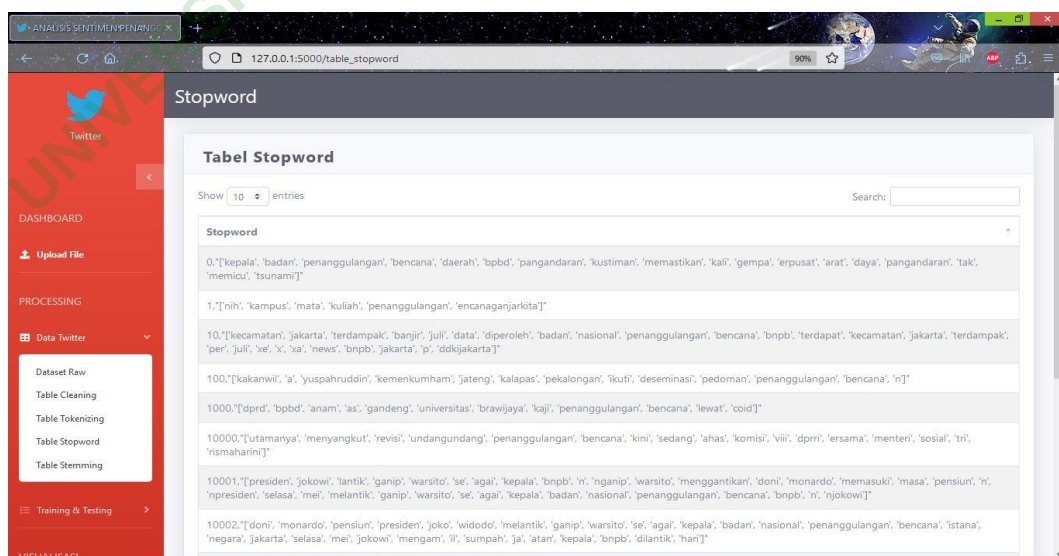
Pada fitur ini data cleaning selanjutnya masuk ke tahap *tokenizing* dari data tersebut akan ada proses pemecahan dari kalimat diubah menjadi beberapa banyak kata. Hasilnya sebagai berikut.



Gambar 4.4 Fitur Table Tokenizing

#### 4.2.5 Fitur Table Stopword

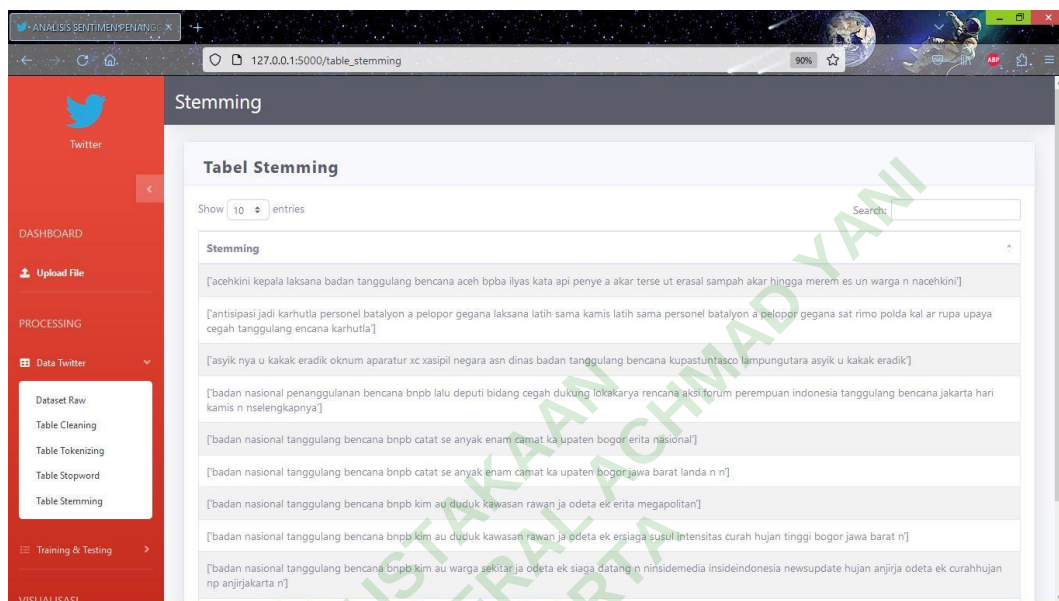
Fitur ini akan diproses untuk menghapus kata yang tidak dibutuhkan pada teks. Hasilnya sebagai berikut.



Gambar 4.5 Fitur Table Stopword

#### 4.2.6 Fitur Table Stemming

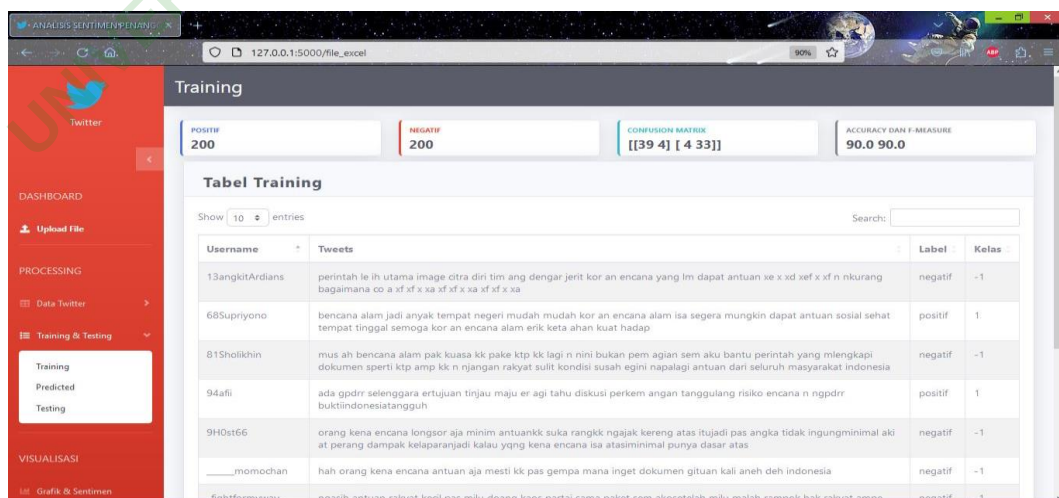
Pada tahap ini akan dilakukan pemetaan dan penguraian dari bentuk suatu kata kemudian diubah dalam bentuk kata yang dasar, Hasil dari stemming sebagai berikut.



Gambar 4.6 Fitur Table Stemming

#### 4.2.7 Fitur Table Training

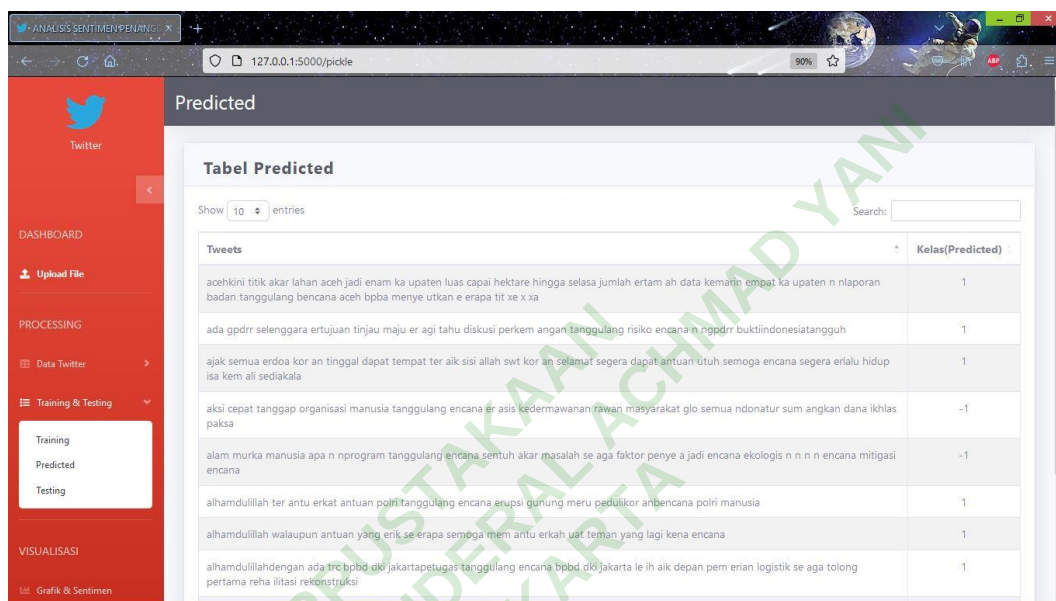
Pada fitur ini berfungsi untuk mencari *accuracy* dan *f-measure* nilai data training yang telah di proses secara otomatis dari sistem. Hasil dari training sebagai berikut.



Gambar 4.7 Fitur Table Training

### 4.2.8 Fitur Predicted

Pada fitur ini adalah proses dari Naïve Bayes Classification yang dimana terdapat model .pickle yang menggunakan data dengan jumlah 400 data yang bagi setiap sentimennya 200 negatif dan 200 positif untuk mendapatkan hasil dari predicted menggunakan sistem. Hasil model tersebut sebagai berikut.



**Tabel Predicted**

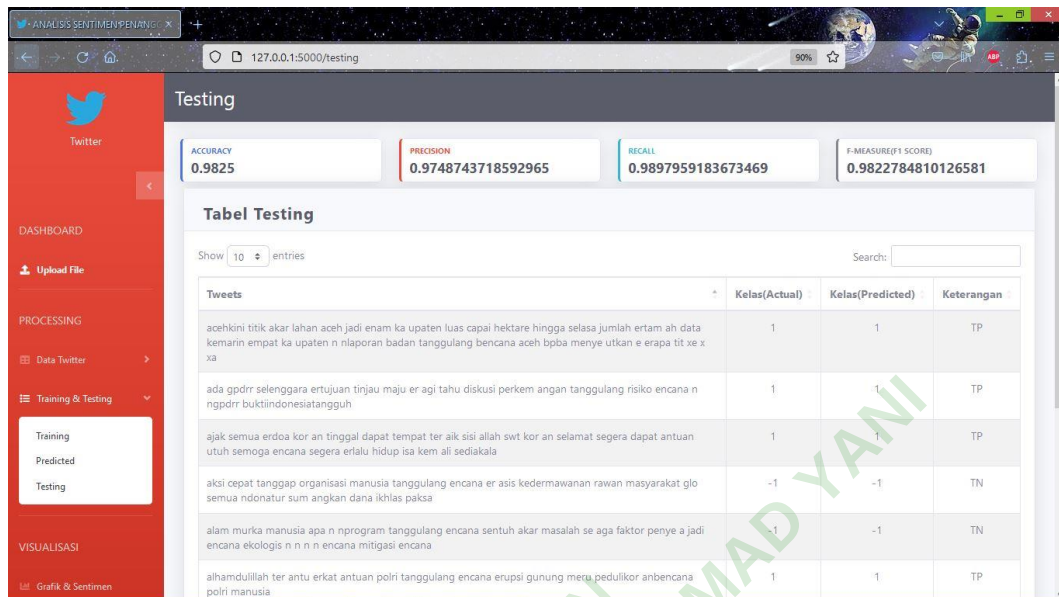
Show 10 entries

| Tweets                                                                                                                                                                                                       | Kelas(Predicted) |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| aceh ini titik akar lahan aceh jadi enam ka upaten luas capai hektare hingga selasa jumlah ertam ah data kemam empat ka upaten n nlaporan badan tanggulang bencana aceh bpba menyenutkan e erapa tit xe x xa | 1                |
| ada gpdr selenggara ertujuan tinjau maju er agi tahu diskusi perkem angan tanggulang risiko encana n ngpdr buktiindonesiatangguh                                                                             | 1                |
| ajak semua erdoa kor an tinggal dapat tempat ter aik sisi allah swt kor an selamat segera dapat antuan utuh semoga encana segera erialu hidup isa kem ali sediakala                                          | 1                |
| aksi cepat tanggap organisasi manusia tanggulang encana er asis kedermawanan rawan masyarakat glo semua ndonatur sum angkan dana ikhlas paksa                                                                | -1               |
| alam murka manusia apa n nprogram tanggulang encana sentuh akar masalah se aga faktor penye a jadi encana ekologis n n n encana mitigasi encana                                                              | -1               |
| alhamdulillah ter antu erikat antuan polri tanggulang encana erupsi gunung meru padulakor an bencana polri manusia                                                                                           | 1                |
| alhamdulillah walaupun antuan yang erik se erapa semoga mem antu erkah uat teman yang lagi kena encana                                                                                                       | 1                |
| alhamdulillah dengan ada tro bpbd di jakartapetugas tanggulang encana bpbd di jakarta le ih aik depan pem erian logistik se aga tolong pertama reha lilitasi rekonstruksi                                    | -1               |

**Gambar 4.8** Fitur Predicted

### 4.2.9 Fitur Testing

Fitur ini fungsinya mencari nilai akurasi testing dengan menggunakan hasil data *actual*/pelabelan dan *predicted*/kelas yang kemudian digabung untuk dapat hasil nilai dari *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Hasilnya tersebut sebagai berikut.



**Testing**

ACCURACY: 0.9825    PRECISION: 0.9748743718592965    RECALL: 0.9897959183673469    F-MEASURE (F1 SCORE): 0.9822784810126581

**Tabel Testing**

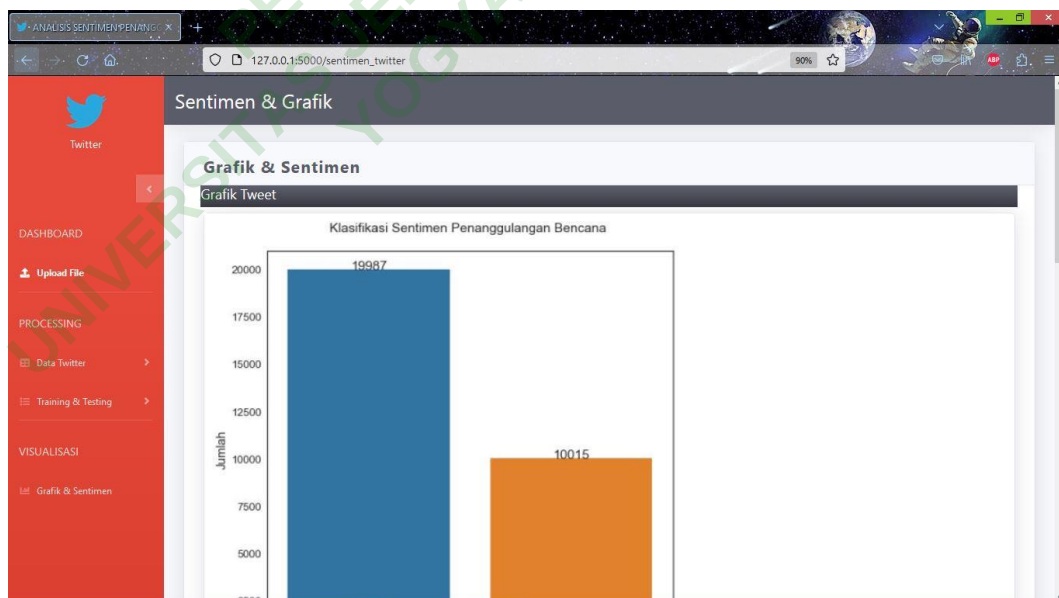
Show 10 entries

| Tweets                                                                                                                                                                                                   | Kelas(Aktual) | Kelas(Predicted) | Keterangan |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|------------------|------------|
| aceh kini titik akar lahan aceh jadi enam ka upaten luas capai hektare hingga selasa jumlah ertam ah data kemarin empat ka upaten n laporan badan tanggulang bencana aceh bpba menyekutkan erapa tit x x | 1             | 1                | TP         |
| ada gpdr selenggara ertujuan tinjau maju er agi tahu diskusi perkem angan tanggulang risiko encana n ngpdr bukti ndonesiatanggul                                                                         | 1             | 1                | TP         |
| ajak semua erdoa kor an tinggal dapat tempat ter aik sisi allah swt kor an selamat segera dapat antuan utuh semoga encana segera erlalu hidup isa kem ali sediakala                                      | 1             | 1                | TP         |
| aksi cepat tanggap organisasi manusia tanggulang encana er asis kedermawanan rawan masyarakat glo semua ndonatur sum angan dana ikhlas paksa                                                             | -1            | -1               | TN         |
| alam murka manusia apa n nprogram tanggulang encana sentuh akar masalah se aga faktor penye a jadi encana ekologis n n n encana mitigasi encana                                                          | -1            | -1               | TN         |
| alhamdulillah ter antu erkat antuan polri tanggulang encana erupsi gunung meru pedulikor an bencana polri manusia                                                                                        | 1             | 1                | TP         |

**Gambar 4.9** Fitur Testing

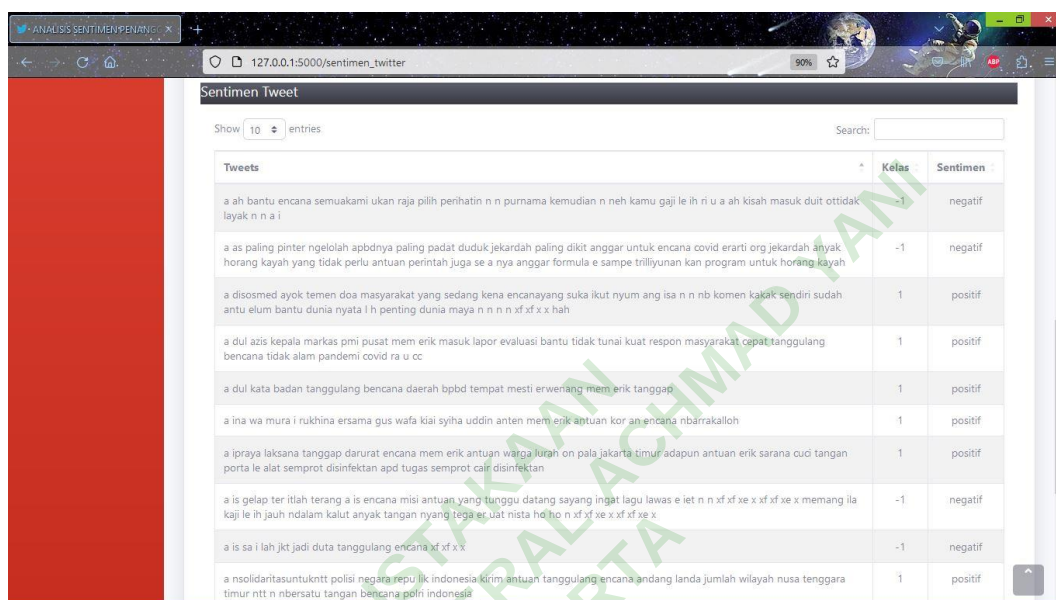
#### 4.2.10 Fitur Grafik & Sentimen

Pada fitur ini hasil dari keseluruhan data berjumlah 32,000 yang telah melalui berbagai tahap awal hingga tahap akhir akan ditampilkan dalam bentuk visualisasi grafik data dan tabel sentimen negatif dan positif. Hasilnya sebagai berikut.



**Gambar 4.10** Fitur Grafik

Dari hasil grafik diatas telah didapatkan bahwa hasil sentimen positif pada grafik berwarna biru dengan jumlah sebanyak 19987 *tweet* dan hasil sentimen negatif pada grafik berwarna oranye dengan jumlah sebanyak 10015 *tweet*, dari grafik tersebut seluruh *tweet* positif dan negatif dapat dilihat pada gambar berikut.



| Tweets                                                                                                                                                                                                                                                        | Kelas | Sentimen |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|----------|
| a ah bantu encana semua kami ukan raja pilih penihatin n n purnama kemudian n neh kamu gaji le ih ri u a ah kisah masuk duit ottidak layak n n a i                                                                                                            | -1    | negatif  |
| a as paling pinter ngelolah apbdnya paling padat duduk jekardah paling dikit anggar untuk encana covid erarti org jekardah anyak horang kayah yang tidak perlu antuan perintah juga se a nya anggar formula e sampe triliyunan kan program untuk horang kayah | -1    | negatif  |
| a disosmed ayok temen doa masyarakat yang sedang kena encanayang suka ikut nyum ang isa n n nb komen kakak sendiri sudah antu elum bantu dunia nyata l h penting dunia maya n n n x f x f x x hah                                                             | 1     | positif  |
| a dul azis kepala markas pmi pusat mem erik masuk lapor evaluasi bantu tidak tunai kuat respon masyarakat cepat tanggulang bencana tidak alam pandemi covid ra u cc                                                                                           | 1     | positif  |
| a dul kata badan tanggulang bencana daerah bpbd tempat mesti erwenang mem erik tanggap                                                                                                                                                                        | 1     | positif  |
| a ina wa mura i rukhina ersama gus wafa koi syiha uddin anten mem erik antuan kor an encana nbarrakalloh                                                                                                                                                      | 1     | positif  |
| a lpraya laksana tanggap darurat encana mem erik antuan warga luah on pala jakarta timur adapun antuan erik sarana cuo tangan porta le alat semprot disinfektan apd tugas semprot calif disinfektan                                                           | 1     | positif  |
| a is gelap ter itlah terang a is encana misi antuan yang tunggu datang sayang ingat lagu lawas e iet n n x f x f x e x x f x f x e x x memang ila kaji le ih jauh ndalam kalut anyak tangan nyang tega er uat nista hdi ho n x f x f x e x x f x f x e x x    | -1    | negatif  |
| a is sa i lah jkt jadi duta tanggulang encana x f x x x                                                                                                                                                                                                       | -1    | negatif  |
| a nsolidantasuntukntt polisi negara republik indonesia kirim antuan tanggulang encana andang landa jumlah wilayah nusa tenggara timur ntt n nbersatu tangan bencana polri indonesia                                                                           | 1     | positif  |

Gambar 4.11 Sentimen

## 4.3 PEMBAHASAN

### 4.3.1 Web Data Extraction

Tahap awal dimulai dengan mengambil data dan mengumpulkan data *tweet* dari Platform Twitter mengenai Penanggulangan Bencana Indonesia dengan sistem Anaconda data tersebut di kelola dalam sistem Jupyter Notebook untuk mendapatkan sekumpulan data dalam bentuk file Microsoft Excel yang akan diproses pada tahap selanjutnya.

### 4.3.2 Kumpulan Data

Pada tahap ini terkumpul data *tweet* mengenai Penanggulangan Bencana menggunakan Anaconda Prompt dengan 2 kata kunci yaitu: Penanggulangan Bencana dan Bantuan Bencana dengan jumlah keseluruhan data sebanyak 30.002, Kemudian menggunakan Jupyter Notebook untuk

dieksekusi sehingga menghasilkan data format CSV yang dapat ditampilkan di Microsoft Excel. Data tersebut merupakan data yang belum dilakukan Processing, bisa diperhatikan pada tabel berikut.

**Tabel 4.1** Kumpulan Data

| No | Datetime,Username,Text                                                                                                                                                                                                                                                                                                                                                                                |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | 2022-07-16 13:46:49+00:00,Jo_arneta0809,"b'@GarudaWisnu11 Boleh juga nih kampus, ada mata kuliah penanggulangan bencana..#GanjarKita"                                                                                                                                                                                                                                                                 |
| 2  | 2022-07-15 03:15:41+00:00,voidotid,"b'Badan Penanggulangan Bencana Daerah ( BPBD) Kabupaten Lebak, Provinsi Banten mengingatkan warga agar waspada hujan lebat disertai petir karena bisa menimbulkan bencana alam.<br><a href="https://t.co/gXJZyPpZ0G">https://t.co/gXJZyPpZ0G</a> "                                                                                                                |
| 3  | 2022-07-13 07:17:50+00:00,BPBDJakarta,"b'Rabu (13/7) Pusdiklat PB BNPB melaksanakan pembukaan acara pelaksanaan kegiatan peningkatan kapasitas petugas penanggulangan bencana BPBD Provinsi DKI Jakarta Angkatan 1 di Gedung Pusdiklat PB BNPB, Sentul Bogor, Jawa Barat.<br>\n\n#KegiatanBPBDJkt #JagaJakarta #SiagaTanggapGalang<br><a href="https://t.co/aFV7T7ABmm">https://t.co/aFV7T7ABmm</a> " |
| 4  | 2022-07-09 16:36:34+00:00,kalap_corona,b'@geloraco Apa ditilep atau dialihkan yang prioritas pada penanggulangan bencana kemanusiaan..dicek yang benaaar..?'                                                                                                                                                                                                                                          |
| 5  | 2022-07-16 12:02:44+00:00,indozonemedia,"b'Pusat Pengendalian Operasi Badan Nasional Penanggulangan Bencana (Pusdalops BNPB) mencatat ada sebanyak 14 kecamatan di Jakarta yang terendam banjir. Namun, berdasarkan kondisi saat ini, genangan air tersebut sudah mulai surut.\n\n <a href="https://t.co/H4zXVfy0OU">https://t.co/H4zXVfy0OU</a> "                                                    |
| 6  | 2022-07-16 11:46:03+00:00,SuaraMabes,b'Samapta Polres Sukabumi cek kesiapan sarana dan prasarana atau peralatan untuk penanggulangan bencana alam <a href="https://t.co/GOT9nApCqH">https://t.co/GOT9nApCqH</a>                                                                                                                                                                                       |
| 7  | 2022-06-29 04:22:20+00:00,pantaugambut,"b'Tapi tetep mesti kita pahami bersama, dalam hal mitigasi bencana karhutla gambut, teknologi ini cuma solusi sementara. Tetep yang paling penting adalah penanggulangan dengan                                                                                                                                                                               |

|    |                                                                                                                                                                                                                                                                                                                                                         |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | upaya restorasi untuk memulihkan gambut yang terlanjur kering supaya basah lagi."                                                                                                                                                                                                                                                                       |
| 8  | 2022-06-27 02:21:32+00:00,Viena_Wu,b'Indonesia memiliki banyak kearifan lokal dalam bidang penanggulangan bencana yang sangat kaya.<br>#LebihSadarBencana <a href="https://t.co/q5RK6GW00M">https://t.co/q5RK6GW00M</a> '                                                                                                                               |
| 9  | 2021-03-08 04:05:00+00:00,kupastuntasnew,"b'Warga yang terdampak tsunami 2018 di Kecamatan Kelumbayan, Kabupaten Tanggamus tidak lama lagi akan memiliki rumah atau hunian tetap (HunTap) yang dibangun Badan Nasional Penanggulangan Bencana (BNPB). <a href="https://t.co/njmP8AVrEF">https://t.co/njmP8AVrEF</a><br>@BNPB_Indonesia @Jurnalist_Tgms" |
| 10 | 2020-02-12 03:13:12+00:00,tempodotco,"b'Pengungsi bencana banjir dan tanah longsor di sejumlah daerah di Lebak, Banten masih membutuhkan banyak bantuan.\n <a href="https://t.co/MkHVgmbwUz">https://t.co/MkHVgmbwUz</a> "                                                                                                                              |
| 11 | 2020-02-10 08:02:06+00:00,DianaCaniago,"b'@taehijaw 2020 masih ada orang gak ada otak nya, mending uangnya buat bantuan orang lain yg ngebutuhin, banyak bencana di awal tahun dan banyak korbannya juga daripada nyebar kebencian kaya gini"                                                                                                           |

#### 4.3.3 Preprocessing Data

Pengolahan data yang bertujuan memperbaiki data belum pernah diperbaiki dari data yang telah tersedia. Sebelum memasuki tahap processing hal yang dilakukan terlebih dahulu yaitu memasukkan kode *library* seperti dibawah ini.

```
from nltk import word_tokenize
from nltk.corpus import stopwords
from werkzeug.utils import secure_filename
import pandas as pd
import numpy as np
import os
import re
import nltk
import pickle
```

Preprocessing data memiliki tahapan sebagai berikut:

## 1. Cleaning text

Pada tahap ini dilakukan proses menghapus atribut yang tidak dibutuhkan berupa url, tag, tanda baca, emoticon, seperti pada kode dibawah ini.

```
def cleaning_text(text):
    tab = text.replace('\t','').replace('\n',' ').replace('\\"',
    ' ')
    score = tab.replace('_', '')
    user = re.sub('@[A-Za-z0-9]+', '', score)
    link = re.sub(
    '((https?):(//)|(\\"\\\"))+([\\w\\d:#@%/;$()~_?\\+-
    =\\\\.\\&] (#!)?)+', '', user)
    url = re.sub(r'http\\S+', '', link)
    punc = re.sub(r'[^\\w\\s]', '', url)
    rt = re.sub(r'RT[\\s]+', '', punc)
    no = re.sub('[0-9]+', '', rt)
    slang = re.sub(r'\\n', " ", no)
    reg = re.sub("b'", " ", slang)
    hashtag = re.sub('/#[\\w_]+[\\t]*/', '', reg)
    result = hashtag.lower()
    emot = emoji.get_emoji_regexp().sub("", hashtag).strip()

    return result
```

Tahap Cleaning telah dilakukan akan menghasilkan sebuah data *tweet* yang bersih, berikut tabel data dari cleaning.

**Tabel 4.2** Data Cleaning

| No | text_clean                                                                                                                                                   |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | badan tanggulang bencana daerah mamuju beri materi mitigasi bencana pelajarbmkg breakingnews opsimediatv                                                     |
| 2  | kalau kampus mata kuliah tanggulang bencanaganjarkita                                                                                                        |
| 3  | badan tanggulang bencana daerah bpbd dki jakarta minta masyarakat waspada potensi rob kira jadi juli beberapa wilayah jakarta utara pulau seribu megapolitan |
| 4  | badan tanggulang bencana daerah bpbd kabupaten lebak catat nilai rugi akibat bencana alam capai rupiah miliar tiga orang lapor tinggal dunia                 |
| 5  | pusat kendali operasi badan nasional tanggulang bencana pusdalops bnpb catat sebanyak camat jakarta rendam berdasarkan kondisi genang air tersebut mulai     |

|    |                                                                                                                                                            |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | surut                                                                                                                                                      |
| 6  | samapta polres sukabumi cek kesiap sarana prasarana alat tanggulang bencana alam                                                                           |
| 7  | badan tanggulang bencana daerah bpbd gunungkidul bersiap antisipasi kering musim kemarau                                                                   |
| 8  | bnpb camat jakarta dampak banjir jakarta dasar lapor terima pusat kendali operasi pusdalops badan nasional tanggulang bencana bnpb banjirjakarta peristiwa |
| 9  | badan tanggulang bencana daerah bpbd dki jakarta catat jumlah genang berita megapolitan                                                                    |
| 10 | badan tanggulang bencana daerah bpbd dki jakarta lapor se anyak jakarta genang hingga sabtu                                                                |
| 11 | camat jakarta dampak banjir juli data oleh badan nasional tanggulang bencana bnpb dapat camat jakarta dampak per juli news bnpb jakarta dkijakarta         |

## 2. Tokenizing

Proses ini dilakukan pemisahan kata dari suatu kalimat kemudian diubah jadi kata tunggal yang dapat berdiri sendiri, kode tokenizing dapat dilihat sebagai berikut.

```
def clean_text(tweet):
    for i in tweet:
        cleaned.append(cleaning_text(
            re.sub("[\n\r\t\xa0]", " ", i).strip()))
    clean_text(data["Text"])
```

## 3. Case folding

Proses mengkonversi seluruh huruf pada kalimat ke huruf yang kecil dalam suatu dokumen, kode *casefolding* dapat dilihat sebagai berikut.

```
def lowercase():
    lower_word = data['text'].str.lower()
    return lower_word

lower_tweet = lowercase()

print(lower_tweet)
```

#### 4. Stopword

Proses untuk hilangkan kata tidak dibutuhkan dan atribut dalam kurun waktu yang sering muncul, berikut ini kode dari Stopword.

```
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory

factory = StopWordRemoverFactory()
more_stopword = ['anjing', 'anjir', 'goblok', 'pantek']
stopword = factory.create_stop_word_remover()
stopwords = factory.get_stop_words()+more_stopword
print(stopwords)
```

Proses stopwords telah dilakukan yang kemudian akan menghasilkan list kata pada tabel berikut ini.

**Tabel 4.3** Kata Stopwords

| Kata                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ['badan', 'penanggulangan', 'bencana', 'daerah', 'bpbdd', 'kota', 'am', 'on', 'mendata', 'se', 'anya', 'k', 'delapan', 'titik', 'lima', 'titik', 'longsor', 'aki', 'at', 'curah', 'hujan', 'tinggi', 'sejak', 'senin', 'juli', 'segala', 'iaya', 'dikeluarkan', 'se', 'agai', 'aki', 'at', 'ditetapkannya', 'keputusan', 'e', 'ankan', 'apbn', 'dana', 'siap', 'pakai', 'badan', 'nasional', 'penanggulangan', 'bencana', 'sum', 'er', 'pem', 'iaayaan', 'lainnya', 'sah', 'mengikat', 'sesuai', 'ketentuan', 'peraturan', 'perundangundangan', 'penyelenggaraan', 'penanganan', 'darurat', 'se', 'agaimana', 'dimaksud', 'diktum', 'kedua', 'dilakukan', 'kemudahan', 'akses', 'se', 'agaimana', 'diatur', 'peraturan', 'perundangundangan', 'mengenai', 'penanggulangan', 'encana', 'pemerintah', 'indonesia', 'melalui', 'badan', 'nasional', 'penanggulangan', 'bencana', 'bnpb', 'menetapkan', 'status', 'keadaan', 'tertentu', 'darurat', 'penyakit', 'mulut', 'kuku', 'pmk', 'hewan', 'ternak', 'melalui', 'surat', 'keputusan', 'kepala', 'bnpb', 'nomor', 'tahun', 'anjing', 'anjir', 'goblok', 'pantek'] |

#### 5. Stemming

Tahap ini mengubah kata untuk membentuk kata dasar perintah. Kode-kode tersebut bisa dilihat sebagai berikut.

```

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

factory = StemmerFactory()
stemmer = factory.create_stemmer()

def stemmed_wrapper(term):
    return stemmer.stem(term)

term_dict = {}

for document in stopwords_tweet:
    for term in document:
        if term not in term_dict:
            term_dict[term] = " "

print(len(term_dict))
print("-----")

for term in term_dict:
    term_dict[term] = stemmed_wrapper(term)
    print(term, ":" ,term_dict[term])

print(term_dict)
print("-----")

def get_stemmed_term(document):
    return [term_dict[term] for term in document]

stem_tweet = stopwords_tweet.apply(get_stemmed_term)

print(stem_tweet)

```

## 6. Normalization

Tahap ini mengubah kata yang disingkat atau salah menjadi sebuah kata yang utuh, kode tersebut bisa dilihat sebagai berikut

```

normalizad_word = pd.read_excel("normalisasi.xlsx")

normalizad_word_dict = {}

for index, row in normalizad_word.iterrows():
    if row[0] not in normalizad_word_dict:
        normalizad_word_dict[row[0]] = row[1]

def normalized_term(document):

```

```

        return [normalizad_word_dict[term] if term in
normalizad_word_dict else term for term in document]

normal_tweet = stem_tweet.apply(normalized_term).str.join("
")

print(normal_tweet)

```

Setelah dilakukan tahap normalisasi akan menghasilkan sebuah data pada tabel berikut ini.

**Tabel 4.4** Kata Normalisasi

| Sebelum    | Sesudah     |
|------------|-------------|
| yucub      | Youtube     |
| Lsung      | Langsung    |
| kpengennya | Kepengennya |
| tidk       | Tidak       |
| spqnduk    | Sepanduk    |
| tingkahnye | Tingkahnya  |
| Ape        | Apa         |
| Bgni       | Begini      |
| Tk         | Tak         |
| koment     | Coment      |
| Bgs        | Bagus       |
| Mjd        | Menjadi     |

#### 4.3.4 Data Clean

Tahap ini merupakan tahap terakhir dari processing, data ini sebagai hasil akhir data yang sudah bersih yang kemudian akan digunakan ke tahap selanjutnya, kode untuk menyimpan data tersebut kedalam bentuk excel dapat dilihat sebagai berikut.

```

clean={"Username":data['Username'],
"Cleaned_Text":data_preprocessing}

```

```
df=pd.DataFrame(clean)
df.to_excel("preprocessing_penanggulangan_bencana_fix.xlsx",
index=False, encoding='utf8')
```

Setelah dilakukan penyimpanan, untuk melihat hasil di atas berikut data tabel yang bisa dilihat.

**Tabel 4.5 Data Clean**

| No | Cleaned_Text                                                                                                                                                   |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | badan tanggulang bencana daerah mamuju beri materi mitigasi bencana pelajarbmkg breakingnews opsimediatv                                                       |
| 2  | kalau kampus mata kuliah tanggulang bencanagankita                                                                                                             |
| 3  | badan tanggulang bencana daerah bpbd dki jakarta minta masyarakat waspada potensi rob kira jadi juli beberapa wilayah jakarta utara pulau seribu megapolitan   |
| 4  | badan tanggulang bencana daerah bpbd kabupaten lebak catat nilai rugi akibat bencana alam capai rupiah miliar tiga orang lapor tinggal dunia                   |
| 5  | pusat kendali operasi badan nasional tanggulang bencana pusdalops bnpb catat sebanyak camat jakarta rendam berdasarkan kondisi genang air tersebut mulai surut |
| 6  | samapta polres sukabumi cek kesiap sarana prasarana alat tanggulang bencana alam                                                                               |
| 7  | badan tanggulang bencana daerah bpbd gunungkidul bersiap antisipasi kering musim kemarau                                                                       |
| 8  | bnpb camat jakarta dampak banjir jakarta dasar lapor terima pusat kendali operasi pusdalops badan nasional tanggulang bencana bnpb banjirjakarta peristiwa     |
| 9  | badan tanggulang bencana daerah bpbd dki jakarta catat jumlah genang berita megapolitan                                                                        |
| 10 | badan tanggulang bencana daerah bpbd dki jakarta lapor se anyak jakarta genang hingga sabtu                                                                    |
| 11 | camat jakarta dampak banjir juli data oleh badan nasional tanggulang bencana bnpb dapat camat jakarta dampak per juli news bnpb jakarta dkijakarta             |

#### 4.3.5 Pelabelan Manual

Suatu proses pelabelan data dari netral, negatif, positif terhadap suatu kata atau kalimat yang dapat analisis kata atau kalimat tersebut mengandung sifat negatif atau positif. Dalam penelitian analisis sentimen ini hanya menggunakan label negatif dan positif, dari data yang telah dilabeli hasilnya sebanyak 400 data *tweet* dengan jumlah masing-masing 200 data negatif dan 200 data positif, berikut hasil dari pelabelan.

|   | username       | cleaned_text                                      | label   | kelas |
|---|----------------|---------------------------------------------------|---------|-------|
| 0 | FauzanMusthofa | tambang pasir ilegal rapi bisa menyebabkan ben... | negatif | -1    |
| 1 | liputan6dotcom | banjir rendam titik kota tangerang badan tangg... | positif | 1     |
| 2 | voidotid       | badan tanggulang bencana daerah bpbd kabupaten... | positif | 1     |
| 3 | BPBDJakarta    | kamis kepala laksana bpbd dki jakarta memberi ... | positif | 1     |
| 4 | mediaindonesia | berbagai potensi dampak kemarau                   | positif | 1     |

**Gambar 4.12** Pelabelan dengan Manual

Hasil dari data pelabelan yang dilakukan dengan manual, berikut data tabelnya.

**Tabel 4.6** Pelabelan Manual

| No | Cleaned_Text                                                                                                                                                                                                                                          | Label   | Kelas |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|-------|
| 1  | badan nasional tanggulang bencana bnpb lalu berbagai program kegiatannya terus laku er aga upaya kurang risiko bencana membangun sadar waspada amp kesiapsiagaan dalam hadap encana untuk meminimalisir jumlah korban akibat encana lebihsadarbencana | positif | 1     |
| 2  | sadar masyarakat jadi kunci tanggulang bencana kalau masyarakat bisa mitigasi upaya tanggulang lain risiko bisa lebih kecil lebihsadarbencana                                                                                                         | positif | 1     |
| 3  | edukasi kaya gini tuh memang kudu sering laku ya biar makin banyak paham soal tanggulang                                                                                                                                                              | positif | 1     |

|    |                                                                                                                                                                                                                                                  |         |    |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|----|
|    | risiko bencana                                                                                                                                                                                                                                   |         |    |
| 4  | badan tanggulang bencana daerah bpbd kota<br>bekas siap tenda dapur umum darurat<br>antisipasi bencana megapolitan<br>jernihkanharapan                                                                                                           | positif | 1  |
| 5  | indonesia landa berbagai bencana gempa<br>banten madura najam passer utara<br>banjirlongsor jayapura erupsi meru jatim<br>sebagian elite bersikeras pindah ibu kota<br>negara bukan lengkap infrastruktur<br>tanggulang bencana tolakibukotabaru | negatif | -1 |
| 6  | badan nasional tanggulang bencana bnpb<br>memberikan bantuan logistik pemkab<br>pandeglang berupa bantuan tahap awal tenda<br>fungsi selimut makan siap saji masker                                                                              | positif | 1  |
| 7  | memang dunia aku bnpb terbentuk badan sar<br>cepat dunia dalam tanggulang bencana cepat<br>tempat bencana                                                                                                                                        | positif | 1  |
| 8  | perlu tanggulang bencana terkait banyak sih<br>mulai hilang materi trauma hilang anggota<br>keluarga lain maka makanya perlu resiliensi<br>pulih tanggulang bencana gpdr<br>buktiindonesiatangguh                                                | positif | 1  |
| 9  | kpk ingat bnpb sektor rawan korupsi<br>tanggulang bencana bnpb<br>bnpbletjensuharyanto firlibahuri korupsi                                                                                                                                       | negatif | -1 |
| 10 | lah memang iya kanpenjajah yang jelas nyata<br>ngapain kamu ngalihin fpi yg lama sering<br>terlibat jadi rawan manusia sertifikat rawan<br>tanggulang bencana tsunami aceh                                                                       | negatif | -1 |

|    |                                                                                                                               |         |    |
|----|-------------------------------------------------------------------------------------------------------------------------------|---------|----|
| 11 | tanggulang bencana tunggang politik<br>olahraga tunggang politik vaksinasi<br>tunggang politik kurang coba yang belum<br>wkwk | negatif | -1 |
|----|-------------------------------------------------------------------------------------------------------------------------------|---------|----|

#### 4.3.6 Training Data

Training data dengan salah satu metode yaitu *Naive Bayes Classification* dengan tahapan awal dari ekstraksi pada teks *Term Frequency-Inverse Document Frequency/TF-IDF* setelah itu proses training untuk melakukan mengklasifikasi sentimen secara otomatis.

Dalam Tabel 3.7 data tersebut akan dilakukan perhitungan pada TF-IDF.

**Tabel 4.7** Dokumen TF-IDF

| Dokumen(d) | Kalimat                                                                                                                                           |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| d1         | kurun waktu tahun indonesia tunjuk progres nyata tanggulang risiko bencana lebih sadar bencana                                                    |
| d2         | badan tanggulang bencana daerah bpbd salur bantuan warga yogyakarta dampak angin puting beliung                                                   |
| d3         | memang dunia akui bnpb terbentuk badan sar cepat dunia dalam tanggulang bencana cepat tempat bencana                                              |
| d4         | apresiasi besar bnpb yang selalu sigap kapan mana tanggulang tanggulang bencana semoga selamat sejahtera selalu serta semua amiin kita lebih siap |

Perhitungan *Term Frequency* (TF) terdapat 6 komponen yang terdiri dari term (kata), d sebagai jumlah dokumen (terdapat 4 dokumen: d1, d2, d3, d4), df sebagai tempat jumlah keseluruhan dari term (kata) yang muncul di dokumen tersedia. Perhitungan *Term Frequency* (TF) dapat dilihat sebagai berikut.

**Tabel 4.8** Perhitungan dari TF

| <b>Term (kata)</b> | <b>d1</b> | <b>d2</b> | <b>d3</b> | <b>d4</b> | <b>df</b> |
|--------------------|-----------|-----------|-----------|-----------|-----------|
| Kurun              | 1         |           |           |           | 1         |
| Waktu              | 1         |           |           |           | 1         |
| Tahun              | 1         |           |           |           | 1         |
| indonesia          | 1         |           |           |           | 1         |
| Tunjuk             | 1         |           |           |           | 1         |
| Progres            | 1         |           |           |           | 1         |
| Nyata              | 1         |           |           |           | 1         |
| tanggulang         | 1         |           |           |           | 1         |
| Risiko             | 1         |           |           |           | 1         |
| bencana            | 2         | 1         | 2         | 1         | 6         |
| Lebih              | 1         |           |           | 1         | 2         |
| Sadar              | 1         |           |           |           | 1         |
| Badan              |           | 1         | 1         |           | 2         |
| tanggulang         |           | 1         | 1         | 2         | 4         |
| Daerah             |           | 1         |           |           | 1         |
| Bpbd               |           | 1         |           |           | 1         |
| Salur              |           | 1         |           |           | 1         |
| bantuan            |           | 1         |           |           | 1         |
| Warga              |           | 1         |           |           | 1         |
| yogyakarta         |           | 1         |           |           | 1         |
| dampak             |           | 1         |           |           | 1         |
| Angin              |           | 1         |           |           | 1         |
| Puting             |           | 1         |           |           | 1         |
| Beliung            |           | 1         |           |           | 1         |
| memang             |           |           | 1         |           | 1         |
| Dunia              |           |           | 2         |           | 2         |
| Akui               |           |           | 1         |           | 1         |
| Bnpb               |           |           | 1         | 1         | 2         |

|           |  |  |   |   |   |
|-----------|--|--|---|---|---|
| terbentuk |  |  | 1 |   | 1 |
| Sar       |  |  | 1 |   | 1 |
| Cepat     |  |  | 2 |   | 2 |
| Dalam     |  |  | 1 |   | 1 |
| Tempat    |  |  | 1 |   | 1 |
| apresiasi |  |  |   | 1 | 1 |
| Besar     |  |  |   | 1 | 1 |
| Yang      |  |  |   | 1 | 1 |
| Selalu    |  |  |   | 2 | 2 |
| Sigap     |  |  |   | 1 | 1 |
| Kapan     |  |  |   | 1 | 1 |
| Mana      |  |  |   | 1 | 1 |
| Semoga    |  |  |   | 1 | 1 |
| selamat   |  |  |   | 1 | 1 |
| sejahtera |  |  |   | 1 | 1 |
| Serta     |  |  |   | 1 | 1 |
| Semua     |  |  |   | 1 | 1 |
| Amiin     |  |  |   | 1 | 1 |
| Kita      |  |  |   | 1 | 1 |
| Siap      |  |  |   | 1 | 1 |

Perhitungan dari IDF(*Inverse Document Frequency*) terdapat beberapa komponen terdiri dari term/kata, df, idf, N. Dari komponen tersebut datanya saling terhubung untuk menentukan hasil dari *Inverse Document Frequency* (IDF). Tabel perhitungannya sebagai berikut.

**Tabel 4.9** Perhitungan dari IDF

| <i>Term</i> (kata) | Df | idf | idf (N=4) | idf (N=1000) |
|--------------------|----|-----|-----------|--------------|
| Kurun              | 1  | 1   | 0.60206   | 3            |
| Waktu              | 1  | 1   | 0.60206   | 3            |

|            |   |         |          |            |
|------------|---|---------|----------|------------|
| Tahun      | 1 | 1       | 0.60206  | 3          |
| indonesia  | 1 | 1       | 0.60206  | 3          |
| Tunjuk     | 1 | 1       | 0.60206  | 3          |
| Progres    | 1 | 1       | 0.60206  | 3          |
| Nyata      | 1 | 1       | 0.60206  | 3          |
| tanggulang | 1 | 1       | 0.60206  | 3          |
| Risiko     | 1 | 1       | 0.60206  | 3          |
| Bencana    | 6 | 0,66666 | 0.176091 | 2.22184875 |
| Lebih      | 2 | 0,5     | 0.30103  | 2.69897    |
| Sadar      | 1 | 1       | 0.60206  | 3          |
| Badan      | 2 | 0,5     | 0.30103  | 2.69897    |
| tanggulang | 4 | 0,25    | 0        | 2.39794001 |
| Daerah     | 1 | 1       | 0.60206  | 3          |
| Bpbd       | 1 | 1       | 0.60206  | 3          |
| Salur      | 1 | 1       | 0.60206  | 3          |
| Bantuan    | 1 | 1       | 0.60206  | 3          |
| Warga      | 1 | 1       | 0.60206  | 3          |
| yogyakarta | 1 | 1       | 0.60206  | 3          |
| Dampak     | 1 | 1       | 0.60206  | 3          |
| Angin      | 1 | 1       | 0.60206  | 3          |
| Puting     | 1 | 1       | 0.60206  | 3          |
| Beliung    | 1 | 1       | 0.60206  | 3          |
| memang     | 1 | 1       | 0.60206  | 3          |
| Dunia      | 2 | 0,5     | 0.30103  | 2.69897    |
| Akui       | 1 | 1       | 0.60206  | 3          |
| Bnpb       | 2 | 0,5     | 0.30103  | 2.69897    |
| terbentuk  | 1 | 1       | 0.60206  | 3          |
| Sar        | 1 | 1       | 0.60206  | 3          |
| Cepat      | 2 | 0,5     | 0.30103  | 2.69897    |
| Dalam      | 1 | 1       | 0.60206  | 3          |

|           |   |     |         |         |
|-----------|---|-----|---------|---------|
| Tempat    | 1 | 1   | 0.60206 | 3       |
| apresiasi | 1 | 1   | 0.60206 | 3       |
| Besar     | 1 | 1   | 0.60206 | 3       |
| Yang      | 1 | 1   | 0.60206 | 3       |
| Selalu    | 2 | 0,5 | 0.30103 | 2.69897 |
| Sigap     | 1 | 1   | 0.60206 | 3       |
| Kapan     | 1 | 1   | 0.60206 | 3       |
| Mana      | 1 | 1   | 0.60206 | 3       |
| Semoga    | 1 | 1   | 0.60206 | 3       |
| Selamat   | 1 | 1   | 0.60206 | 3       |
| sejahtera | 1 | 1   | 0.60206 | 3       |
| Serta     | 1 | 1   | 0.60206 | 3       |
| Semua     | 1 | 1   | 0.60206 | 3       |
| Amiin     | 1 | 1   | 0.60206 | 3       |
| Kita      | 1 | 1   | 0.60206 | 3       |
| Siap      | 1 | 1   | 0.60206 | 3       |

Pada penelitian ini proses Klasifikasi mengekstraksi TF-IDF sehingga data tersebut secara otomatis mendapatkan hasil dari nilai kata di salah satu dokumen, Perhitungan tersebut memiliki *library* `sklearn.feature_extraction.text` & `TfidfVectorizer` yang dapat menjalankan otomatis data tersebut dan terdapat *library* Multinomial Naïve Bayes untuk klasifikasi data training berupa teks data. Berikut perhitungannya.

**Tabel 4.10** Perhitungan dari TF-IDF

| <b>Term (kata)</b> | <b>d1</b> | <b>d2</b> | <b>d3</b> | <b>d4</b> |
|--------------------|-----------|-----------|-----------|-----------|
| kurun              | 0.60206   |           |           |           |
| waktu              | 0.60206   |           |           |           |
| tahun              | 0.60206   |           |           |           |
| indonesia          | 0.60206   |           |           |           |

|            |          |          |          |          |
|------------|----------|----------|----------|----------|
| tunjuk     | 0.60206  |          |          |          |
| progres    | 0.60206  |          |          |          |
| nyata      | 0.60206  |          |          |          |
| tanggulang | 0.60206  |          |          |          |
| risiko     | 0.60206  |          |          |          |
| bencana    | 0.176091 | 0.176091 | 0.176091 | 0.176091 |
| lebih      | 0.30103  |          |          | 0.30103  |
| sadar      | 0.60206  |          |          |          |
| badan      |          | 0.30103  | 0.30103  |          |
| tanggulang |          | 0        | 0        | 0        |
| daerah     |          | 0.60206  |          |          |
| bpbd       |          | 0.60206  |          |          |
| salur      |          | 0.60206  |          |          |
| bantuan    |          | 0.60206  |          |          |
| warga      |          | 0.60206  |          |          |
| yogyakarta |          | 0.60206  |          |          |
| dampak     |          | 0.60206  |          |          |
| angin      |          | 0.60206  |          |          |
| puting     |          | 0.60206  |          |          |
| beliung    |          | 0.60206  |          |          |
| memang     |          |          | 0.60206  |          |
| dunia      |          |          | 0.30103  |          |
| akui       |          |          | 0.60206  |          |
| bnpb       |          |          | 0.30103  | 0.30103  |
| terbentuk  |          |          | 0.60206  |          |
| sar        |          |          | 0.60206  |          |
| cepat      |          |          | 0.30103  |          |
| dalam      |          |          | 0.60206  |          |
| tempat     |          |          | 0.60206  |          |
| apresiasi  |          |          |          | 0.60206  |

|           |  |  |  |         |
|-----------|--|--|--|---------|
| besar     |  |  |  | 0.60206 |
| yang      |  |  |  | 0.60206 |
| selalu    |  |  |  | 0.30103 |
| sigap     |  |  |  | 0.60206 |
| kaplan    |  |  |  | 0.60206 |
| mana      |  |  |  | 0.60206 |
| semoga    |  |  |  | 0.60206 |
| selamat   |  |  |  | 0.60206 |
| sejahtera |  |  |  | 0.60206 |
| serta     |  |  |  | 0.60206 |
| semua     |  |  |  | 0.60206 |
| amiin     |  |  |  | 0.60206 |
| kita      |  |  |  | 0.60206 |
| siap      |  |  |  | 0.60206 |

Dari perhitungan sistem di atas terdapat kode sebagai berikut.

```
s1 = "semua badan bagus bahkan badan penanggulangan bencana daerah
sekalipun"
s2 = "negara sedang darurat darurat ketidakbecusan penanggulangan
bencana "

vect = TfidfVectorizer()
X = vect.fit_transform([s1, s2])

X.toarray()
```

```
[(0.632060998172729, 'badan'),
 (0.3160304990863645, 'bagus'),
 (0.3160304990863645, 'bahkan'),
 (0.22485830296545706, 'bencana'),
 (0.3160304990863645, 'daerah'),
 (0.0, 'darurat'),
 (0.0, 'ketidakbecusan'),
 (0.0, 'negara'),
 (0.22485830296545706, 'penanggulangan'),
 (0.0, 'sedang'),
 (0.3160304990863645, 'sekalipun'),
 (0.3160304990863645, 'semua')]
```

**Gambar 4.13** Hasil Pertama Perhitungan TF-IDF

```
[(0.0, 'badan'),
 (0.0, 'bagus'),
 (0.0, 'bahkan'),
 (0.2513600402461016, 'bencana'),
 (0.0, 'daerah'),
 (0.7065555322771194, 'darurat'),
 (0.3532777661385597, 'ketidakbecusan'),
 (0.3532777661385597, 'negara'),
 (0.2513600402461016, 'penanggulangan'),
 (0.3532777661385597, 'sedang'),
 (0.0, 'sekalipun'),
 (0.0, 'semua')]
```

**Gambar 4.14** Hasil Kedua Perhitungan TF-IDF

Setelah perhitungan TF-IDF selesai dilakukan kemudian mencari nilai akurasi data training untuk mendapatkan keakuratan data tersebut menggunakan data yang sudah dilakukan pelabelan manual. Kode dari akurasi data bisa dilihat sebagai berikut.

```
print("Accuracy: {:.2f}%".format(accuracy_score(y_test, y_pred) *
100))
print("\nF1 Score: {:.2f}".format(f1_score(y_test, y_pred,
average='weighted') * 100))
print("\nConfusion Matrix:\n", confusion_matrix(y_test, y_pred))
```

Cross-validation yaitu suatu metode yang berguna dalam mencari akurasi data, prinsip yang dimiliki Cross-validation yaitu membagi 2 bagian data yang terdiri dari data latih dan data uji, penggunaan *library* from `sklearn.model_selection` dan import `ShuffleSplit` untuk melakukan perhitungan dengan rata-rata nilai 10 kali. Kode dari perhitungan nilai cross validation sebagai berikut.

```
fig, (ax1, ax2) = plt.subplots(2, 1, sharex=True,
figsize=(16,9))

acc_scores = [round(a * 100, 1) for a in accs]
f1_scores = [round(f * 100, 2) for f in f1s]
```

```

x1 = np.arange(len(acc_scores))
x2 = np.arange(len(f1_scores))

ax1.bar(x1, acc_scores)
ax2.bar(x2, f1_scores, color='#559ebf')

# Place values on top of bars
for i, v in enumerate(list(zip(acc_scores, f1_scores))):
    ax1.text(i - 0.25, v[0] + 2, str(v[0]) + '%')
    ax2.text(i - 0.25, v[1] + 2, str(v[1]))

ax1.set_ylabel('Accuracy (%)')
ax1.set_title('Naive Bayes')
ax1.set_ylim([0, 100])

ax2.set_ylabel('F1 Score')
ax2.set_xlabel('Runs')
ax2.set_ylim([0, 100])

sns.despine(bottom=True, left=True)
plt.show()

```

Proses perhitungan dilanjutkan tahap menghitung keakuratan model dari tahap training sehingga dapat prediksi kelas/label dari data tersedia. Kemudian langkah selanjutnya membuat model klasifikasi X dan Y sebagai variabel, dengan adanya pembuatan model ini bertujuan untuk mempermudah data tersebut di proses ke tahap selanjutnya.

Pemodelan untuk klasifikasi membutuhkan *library* sklearn.pipeline dan *import* pipeline fungsinya yaitu *testable* saat melakukan cross-validation, *import* pickle fungsinya yaitu membaca dan menyimpan data dengan format .pkl. *import library* telah dilakukan langkah selanjutnya yaitu membuat model klasifikasi dengan menggunakan kode sebagai berikut.

```

X = df.Komentar
y = df.Klasifikasi

text_classifier = Pipeline([('vect', TfidfVectorizer()),
                             ('tfidf', TfidfTransformer()),
                             ('classifier', MultinomialNB(alpha=1.0)),
                             ])
X_train = np.asarray(X)
text_classifier = text_classifier.fit(X_train, np.asarray(y))

```

Model variabel yang telah diberi nama *text\_classifier* akan disimpan dengan format .pickle, untuk kodenya dapat dilihat sebagai berikut.

```
files = open('model_nbc.pickle', 'wb')
pickle.dump(text_classifier, files)
files.close()
```

File *pickle* dapat dibuka kembali dengan cara memasukkan kode seperti berikut ini.

```
model = open('model_nbc.pickle', 'rb')
nbc_classifier = pickle.load(model)
print(nbc_classifier)
```

Kemudian data yang telah dilakukan klasifikasi kemudian di proses menggunakan data testing, terdapat 200 *tweet* data telah dilakukan pelabelan dengan cara manual yang berjumlah 400 data dari training, serta 200 data *tweet* dari jumlah 32000 data *tweet* yang digunakan untuk testing.

Langkah berikutnya yaitu mencari hasil prediksi Naïve Bayes, dengan cara memasukkan kode sebagai berikut.

```
predicted = nbc_classifier.predict(np.asarray(data_tweet))
predicted
```

Setelah itu data hasil prediksi akan disimpan dalam variabel `result_tweet`, Untuk dapat menyimpan data tersebut yaitu dengan memasukkan kode sebagai berikut.

```
result_tweet=[]
for i in range(len(predicted)):
    if(predicted[i]==1):
        sentiment_result='positif'
    elif(predicted[i]==-1):
        sentiment_result='negatif'
    # result_tweet.append({'class':prediction_linear[i],
    'result_nbc':sentiment_result})
    result_tweet.append({'cleaned_text':data_tweet[i],
    'class':predicted[i] })
```

#### 4.3.7 Testing

Pada tahapan ini untuk melihat keakuratan dari pemodelan yang dilakukan dalam memprediksi label kelas pada data uji.

|    | Text                                                | Actual | Predicted |
|----|-----------------------------------------------------|--------|-----------|
| 0  | tambang pasir ilegal merapi bisa menyebabkan b...   | -1     | -1        |
| 1  | banjir rendam titik kota tangerang badan tangg...   | 1      | 1         |
| 2  | badan tanggulang bencana daerah bpbd kabupaten...   | 1      | 1         |
| 3  | kamis kepala pelaksana bpbd dki jakarta member...   | 1      | 1         |
| 4  | berbagai potensi dampak kemarau kota sukabumi ...   | 1      | 1         |
| 5  | meski tahun musim kemarau relatif basah badan ...   | 1      | 1         |
| 6  | badan nasional penanggulangan bencana bnpb lal...   | 1      | 1         |
| 7  | halo sahabattangguh saat seluruh provinsi indo...   | 1      | 1         |
| 8  | aksi cepat tanggap organisasi manusia tanggula...   | -1     | -1        |
| 9  | rabu bpbd dki jakarta melaksanakan giat pembinaa... | 1      | 1         |
| 10 | rabu tanggal pukul serti didik purwanto melaks...   | 1      | 1         |
| 11 | badan penanggulangan bencana daerah bpbd kabup...   | 1      | 1         |
| 12 | selasa bpbd diy bersama forum prb diy laksanak...   | 1      | 1         |
| 13 | bpbd kota ambon lmelalui pusat kendali operasi...   | 1      | 1         |
| 14 | konsep peer to peer memberdayakan difabel sama...   | 1      | 1         |
| 15 | perintah lalu badan nasional penanggulangan be...   | 1      | 1         |
| 16 | syukur pasti indonesia capai berbagai progress...   | 1      | 1         |
| 17 | wujud resiliensi berkelanjutan indonesia milik...   | 1      | 1         |
| 18 | kurun waktu tahun indonesia tunjuk progres nya...   | 1      | 1         |
| 19 | masyarakat peduli lingkungan tidak membuang sa...   | 1      | 1         |

**Gambar 4.15** Hasil pelabelan Manual dan Naïve Bayes

#### 4.3.8 Hasil Klasifikasi

Hasil klasifikasi didapatkan melalui proses mencari nilai akurasi dari tahap training dan testing yang kemudian dilakukan tahap klasifikasi untuk mendapatkan hasil data yang mengandung sentimen negatif serta positif. Hasil dari data tersebut sebagai berikut.

|    | cleaned_text                                      | klasifikasi |
|----|---------------------------------------------------|-------------|
| 0  | kepala badan tanggulang bencana daerah bpbd pa... | 1           |
| 1  | kalau kampus mata kuliah tanggulang encanaganj... | -1          |
| 2  | video jakarta rendam banjir tinggi capai cm ke... | 1           |
| 3  | garut kota tarogong kidul n npemda garut kerah... | 1           |
| 4  | pusat kendali operasi badan nasional tanggulan... | 1           |
| 5  | samapta polres suka umi cek kesiap sarana pras... | 1           |
| 6  | badan tanggulang bencana daerah dki jakarta me... | 1           |
| 7  | bnpb camat jakarta dampak banjir n njakarta xe... | 1           |
| 8  | badan tanggulang bencana daerah bpbd dki jakar... | 1           |
| 9  | badan tanggulang bencana daerah bpbd dki jakar... | 1           |
| 10 | camat jakarta dampak banjir juli data oleh bad... | 1           |
| 11 | badan nasional tanggulang bencana bnpb kim au ... | 1           |
| 12 | tingkat profesionalisme lpbi nu ka upaten mala... | 1           |

**Gambar 4.16** Hasil Data Sentimen Positif dan Negatif

Dari evaluasi model data klasifikasi yang digunakan berjumlah 400 data dengan masing-masing 200 data negatif dan 200 data positif yang sudah dilakukan pelabelan manual. Langkah selanjutnya yaitu melakukan pengujian pada *confusion matrix* untuk mengetahui nilai dari data aktual dan data prediksi sekaligus mengecek aplikasi berjalan dengan baik.

Kemudian dilakukan perhitungan *cross validation* menggunakan *confusion matrix* dalam menguji akurat melalui nilai dari data aktual dan data prediksi. Hasil dari perhitungan *cross validation* sebagai berikut.

**Tabel 4.11** Confusion Matrix dari Data Training

| Kelas Prediksi | Kelas Aktual |         |
|----------------|--------------|---------|
|                | Positif      | Negatif |
| Positif        | 187          | 13      |
| Negatif        | 8            | 192     |

Dari data diatas maka didapatkan hasil yaitu  $TP = 187$ ,  $TN = 13$ ,  $FP = 8$ ,  $FN = 192$ . Proses *confusion matrix* telah selesai kemudian dilanjutkan tahap menguji data k-fold cross validation dalam mencari angka sesuai yang dengan 10 kali perulangan data. Dari setiap fold terdapat k-fold cross validation yang telah diketahui nilai dan hasil berbeda. Hasil nilainya sebagai berikut.

**Tabel 4.12** Nilai dari K-fold Accuracy dan F-measure

| Fold    | Accuracy | F-Measure |
|---------|----------|-----------|
| Fold 1  | 93,8%    | 93,75%    |
| Fold 2  | 91,2%    | 91,24%    |
| Fold 3  | 90,0%    | 90,04%    |
| Fold 4  | 88,8%    | 88,73%    |
| Fold 5  | 97,5%    | 97,51%    |
| Fold 6  | 95,0%    | 95,0%     |
| Fold 7  | 90,0%    | 90,03%    |
| Fold 8  | 88,8%    | 88,79%    |
| Fold 9  | 88,8%    | 88,59%    |
| Fold 10 | 96,2%    | 96,25%    |

Dapat dilihat Tabel diatas bahwa nilai rata – rata yang dihasilkan untuk akurasi yaitu nilai 91 *accuracy* dan nilai 91 *f-measure*.

Proses selanjutnya masuk ke tahap *testing dari data* dengan menggunakan 400 *tweet* data yang telah dilakukan pelabelan manual 200 negatif dan 200 positif. Dalam mencari nilai akurasi dari testing tersebut dibutuhkan *confusion matrix* mendapatkan perbedaan antara nilai data dari *training* dan nilai data dari *testing*. Dari perhitungan tersebut bisa dilihat hasilnya sebagai berikut.

**Tabel 4.13** Confusion Matrix dari Data Testing

| Kelas Prediksi | Kelas Aktual |         |
|----------------|--------------|---------|
|                | Positif      | Negatif |
| Positif        | 194          | 6       |
| Negatif        | 5            | 195     |

Dari diatas telah diketahui  $TP = 194$ ,  $TN = 199$ ,  $FP = 6$ ,  $FN = 5$  menggunakan persamaan (1) untuk mengetahui nilai akurasi dan rumus persamaan (5) untuk mengetahui nilai f-measure. Hasil tersebut dapat dibawah ini.

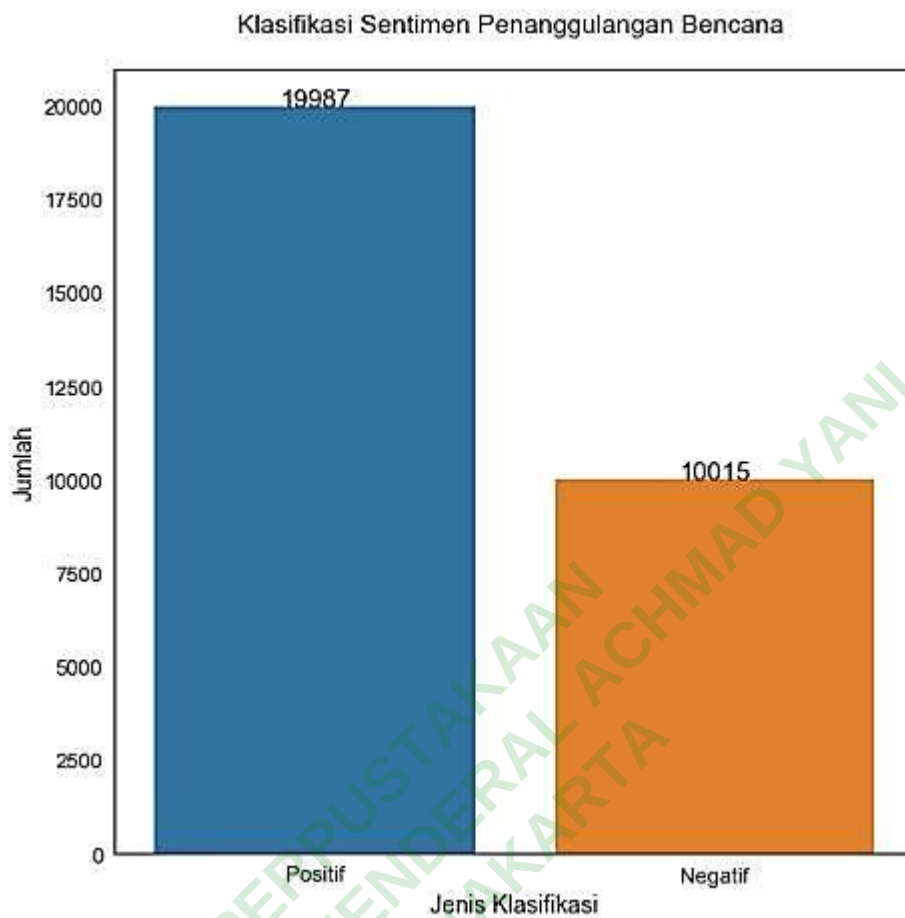
**Tabel 4.14** Akurasi dari Data Testing

| <b>Akurasi Data Testing</b> | <b>Hasil</b> |
|-----------------------------|--------------|
| <b>Accuracy</b>             | 98,25%       |
| <b>F-measure</b>            | 98,22%       |

Dari data di atas diketahui bahwa nilai dari data *accuracy* dan data *f-measure* pada data dari testing dengan jumlah 400 tweet yang telah dilabeli manual menghasilkan nilai *accuracy* sebanyak 98,25% dan *f-measure* sebanyak 98,22% sehingga menghasilkan perbedaan 0,03 % pada data tersebut.

#### **4.4 HASIL ANALISIS**

Pada tahap ini adalah proses menghitung total keseluruhan data *Twitter* yang telah diproses sebanyak 32.000 terdiri dari data 400 yang telah dilabeli manual, pada tahap training sebanyak 31.200 data yang kemudian data 400 proses untuk proses testing, Hasil dari data yang telah diprediksi baik positif dan negatif secara keseluruhan sebagai berikut.



**Gambar 4.17** Histogram pada data Sentimen Positif dan Negatif

Dari data di atas telah diketahui bahwa data keseluruhan didapatkan sentimen dari positif sebanyak 19,987 yang di dalam data tersebut para pengguna membahas berbagai topik terkait baiknya kinerja badan penanggulangan bencana terhadap penanggulangan bencana alam, cepatnya bantuan datang saat terjadinya bencana alam, proses pemulihan wilayah bencana yang baik, harapan dan dukungan terhadap penanggulangan bencana alam. Dari hasil diatas terdapat contoh data *tweet* sentimen positif sebagai berikut.

**Tabel 4.15** Data Tweet Sentimen Positif

| No | Data Tweet                                                                                                                                        |
|----|---------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | maka menghadapi bencana alam perintah berkolaborasi banyak pihak masuk dorong investasi sains punya infrastruktur canggih tanggulung bencana gpdr |

|   |                                                                                                                                                                            |
|---|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|   | bukti Indonesia tangguh                                                                                                                                                    |
| 2 | maka betul negara pernah kena bencana kalau tempat sharing cara tangani tanggulang lain bakal oke sih                                                                      |
| 3 | semua badan bagus bahkan badan tanggulang bencana daerah sekalipun                                                                                                         |
| 4 | makanya enggak heran kalo bnpb sangat giat ajak melibatkan masyarakat masuk dunia usaha akademisi media untuk edukasi sosialisasi cegah tanggulang bencana kita lebih siap |
| 5 | semangat terus bnpb tanggulang bencana indonesia yang sistematis padu yukkk kita lebih siap                                                                                |

Sedangkan pada sentimen negatif sebanyak 10,015 yang di dalam data tersebut para pengguna membahas berbagai topik terkait lambatnya penyaluran bantuan bencana alam, proses penanggulangan bencana yang tidak cepat dan tanggap, proses pengambilan bantuan yang aturannya rumit, kinerja badan penanggulangan bencana yang tidak maksimal. Dari hasil tersebut terdapat contoh data *tweet* sentimen negatif sebagai berikut.

**Tabel 4.16** Data dari Sentimen Tweet Negatif

| No | Data Tweet                                                                                                                                                     |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | negara sedang darurat darurat tidak becusan tanggulang bencana                                                                                                 |
| 2  | gilir korban bencana suka ambil bantuan suruh bawa kk padahal rumah hancur le atau gelandang langsung angkat jadi pegawai bumh jangan tanya kk ktp aja mungkin |
| 3  | maksud gila beneran gila kena bencana butuh bantuan kudu pake kk kali aja sempet masuk rumah rumah sudah rata                                                  |
| 4  | bukan mikir tanggulang bencana sama pandemi malah sibuk pindah another level memang                                                                            |
| 5  | henti berbagai giat tidak perlu tanggulang bencana tidak usah pejabat berbondong bondong datang lokasi bencana                                                 |

Dari hasil data yang telah dijelaskan dapat disimpulkan bahwa banyak para pengguna Twitter yang memberikan *tweet* sentimen positif sebanyak 19,987 dalam menanggapi berbagai topik kinerja badan penanggulangan, bantuan bencana, pemulihan wilayah bencana. Sehingga penanggulangan bencana di indonesia ini masih termasuk dalam kategori yang baik.