

BAB 3

METODE PENELITIAN

Penelitian ini merupakan penelitian analisis sentiment terkait *tweet* (cuitan) masyarakat pada Twitter dengan nilai positif dan negatif menggunakan metode *Support Vector Machine*. Dalam penelitian ini membutuhkan data dari *tweet* (cuitan) yang diperoleh dari Twitter terkait penolakan yang rencana proyek penambangan batu andesit yang dilaksanakan di desa Wadas, Kecamatan Bener, Kabupaten Purworejo, Jawa Tengah. Kemudian dilakukan pengolahan data *tweet* berupa *preprocessing* untuk mendapatkan kumpulan data yang diinginkan. Data ini kemudian akan digunakan untuk mengkategorikan informasi atau persepsi pengguna internet di Twitter mengenai konflik yang terjadi di dalam Wadas sehingga terinformasikan sebagaimana mestinya.

Penelitian ini diambil dari latar belakang permasalahan yang ada, mengolah data yang diperoleh sehingga dapat ditentukan sentimen yang benar berdasarkan informasi yang diperoleh. Di bawah ini adalah bahan penelitian, *tools*, dan *workflow* analisis sentimen Wadas Purworejo beserta langkah-langkah penelitian untuk melengkapi analisis sentimen menggunakan data *tweet*.

Beberapa metode untuk mendapatkan data atau informasi dalam melakukan pemecahan masalah. Berikut adalah beberapa metode yang akan digunakan, antara lain:

1. Studi Pustaka

Metode pengumpulan data menggunakan metode kepastakaan ini meliputi pengumpulan jurnal, *literatur*, dokumen, artikel maupun laman web sebagai sumber pustaka yang berhubungan dengan materi khususnya analisis sentimen menggunakan metode *Support Vector Machine (SVM)*.

2. Pengumpulan Data *Tweet* (Cuitan).

Data yang diperoleh adalah data yang didapatkan secara langsung dari Twitter dengan menggunakan *jupyter notebook* terkait dengan

pencarian #SaveWadas #WadasMelawan #WadasTolakTambang #CabutPLWadas.

3. Perancangan.

Dalam proses ini, beberapa hal yang akan dilakukan termasuk desain alur pengambilan data, metode sistem dan pemrograman.

4. Pengujian dan Evaluasi.

Pada tahap ini, pengujian pengklasifikasi *Support Vector Machine* dilakukan untuk menemukan kesalahan yang mungkin terjadi dan dapat diperbaiki.

5. Penyusunan laporan

Pada tahap ini, penyusunan laporan dilakukan sebagai dokumentasi serta memfasilitasi pembelajaran dan pengembangan orang lain.

3.1 BAHAN PENELITIAN

Bahan dari penelitian ini menjelaskan dari data *tweet* (cuitan) yang ada di dalam Twitter terkait dengan penolakan proyek penambangan batu andesit yang dilakukan di Desa Wadas, Kecamatan Bener, Kabupaten Purworejo, Jawa Tengah terkait hastag #SaveWadas #WadasMelawan #WadasTolakTambang #CabutPLWadas.

3.2 ALAT PENELITIAN

Alat yang dipergunakan pada penelitian ini yaitu dengan laptop spesifikasi yang baik serta dapat melakukan pengolahan data serta terhubung dengan jaringan internet.

Sistem Operasi serta program aplikasi yang dipergunakan pada pengembangan aplikasi ini adalah:

1. Sistem Operasi: Windows 8 atau lebih baru.
2. Jupyter Notebook Versi 3.1
3. Microsoft Excel 2019.
4. Anaconda Versi 3.
5. Bahasa Pemrograman Python 3.7.4.
6. Sublime Text 3

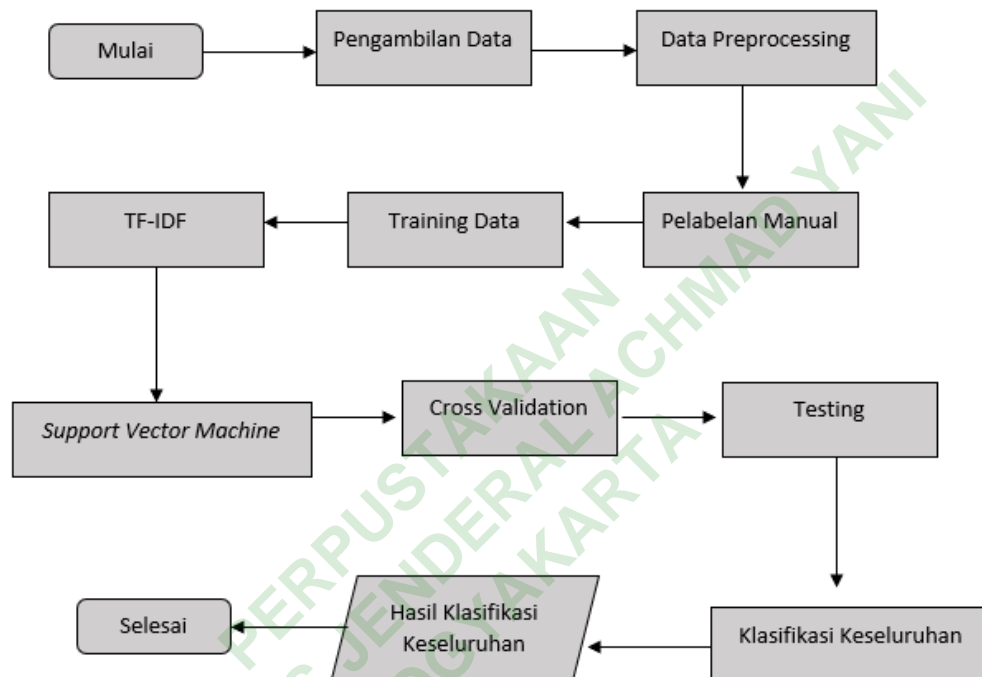
Library yang digunakan dalam penelitian ini sebagai berikut:

- a. *Numpy* : *library* ini dipergunakan dalam pembuatan objek kelas array tunggal kedalam multidimensi
- b. *Pandas* : *library* ini dipergunakan untuk mengatur letak data sehingga mudah untuk dicari secara intuitif
- c. *Nltk* : *library* ini dipergunakan dalam mempermudah dalam pemrosesan data. Contoh pemrosesan data seperti melakukan *classification*, *tokenization*, *stemming*, *tagging* dan *parsing*.
- d. *Sastrawi* : *library* ini dipergunakan untuk mengubah kata berimbuhan bahasa Indonesia menjadi bentuk dasarnya.
- e. *Sklearn* : *library* ini dipergunakan untuk membantu dalam melakukan *processing* data dan melakukan *training* data untuk kebutuhan *machine learning*.
- f. *Pickle* : *library* ini dipergunakan untuk menyimpan dan memanggil suatu file berformat .pkl.
- g. *Snsrape* : *library* ini dipergunakan untuk pengambilan data Twitter tanpa menggunakan API.
- h. *Seabron* : *library* ini dipergunakan untuk menentukan hubungan antara dua variabel
- i. *Matplotlib* : *library* ini dipergunakan untuk memvisualisasikan data dengan menggunakan *Numpy*.

3.3 JALAN PENELITIAN

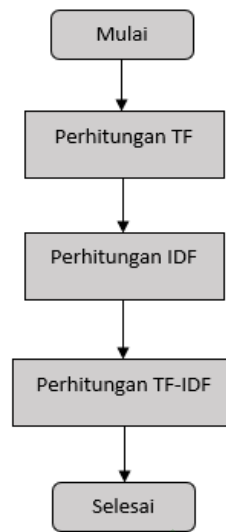
Penelitian ini adalah penelitian analisis sentimen mengenai *tweet* (cuitan) warganet dari Twitter dengan nilai positif dan negatif dengan menggunakan metode *Support Vector Machine*. Penelitian ini membutuhkan data *tweet* (cuitan) yang diperoleh dari Twitter mengenai penolakan rencana proyek penambangan batu andesit yang dilakukan di desa Wadas, Kecamatan Bener, Kabupaten Purworejo, Jawa Tengah. Selain itu, dilakukan pengolahan dataset berupa preprocessing untuk mendapatkan dataset yang sesuai. Data *preprocessing* merupakan Teknik data mining yang umum digunakan untuk mengubah data mentah menjadi format yang

dapat dimengerti. Dalam penelitian ini bahasa pemrograman yang digunakan adalah *Python*, *Jupyter Notebook* dan *Anaconda 3* yang digunakan dalam pengambilan *tweet* data yang akan ditampilkan pada *Microsoft Office Excel* serta akan dilakukan pemodelan dengan *library*. Pada Gambar 3.1 merupakan jalan penelitian yang akan dilakukan.



Gambar 3.1 Flowchart Jalan Penelitian

Pada tahapan *training* akan dilakukan beberapa proses TF-IDF dapat dilihat dalam Gambar 3.2.



Gambar 3.2 Flowchart TF-IDF

Berikut merupakan langkah-langkah jalannya penelitian yang akan dilakukan, yaitu:

3.3.1 Pengambilan Data.

Pada tahap ini, akan diambil data dari Twitter untuk mendapatkan data *tweet* mengenai konflik yang terjadi di Wadas menggunakan perangkat lunak *Anaconda Prompt* kemudian akan dijalankan pada *Jupyter Notebook* sehingga akan mendapatkan data dengan format *csv* yang ditampilkan kedalam *Microsoft Excel*. Data *tweet* yang diambil merupakan *tweet* mengenai *hashtag* “#SaveWadas #WadasMelawan #WadasTolakTambang #CabutPLWadas” selama periode 01 Januari 2021 – 22 Juli 2022 dengan mendapatkan jumlah *tweet* 13397. Dalam pengambilan data ini akan digunakan beberapa *library* yang ada di *Python*. *Snsrape* merupakan *library* yang dipergunakan. *Library snsrape* digunakan karena tidak memerlukan API Twitter sehingga tidak perlu mengakses akun Developer Twitter. Berikut merupakan kode untuk mengimport *library* pengambilan data dapat dilihat dalam Gambar 3.3.

```
import snsrape.modules.twitter as sntwitter
import pandas as pd
```

Gambar 3.3 Library Pengambilan Data

Tweet yang didapat pada pengambilan data ini berformat *csv* yang kemudian akan dimanipulasi dengan menggunakan *pandas* untuk pengambilan *text* dan *datetime*. Kode yang dipergunakan dalam mengambil *tweet* dapat dilihat dalam Gambar 3.4.

```
# Setting variables to be used below
maxTweets = 10000

# Creating list to append tweet data to
tweets_list2 = []

# Using TwitterSearchScrapper to scrape data and append tweets to list
for i,tweet in enumerate(sntwitter.TwitterSearchScrapper('#cabutPLWadas since:2021-01-01 until:2022-07-22 lang:id').get_items()):
    if i>maxTweets:
        break
    tweets_list2.append([tweet.date, tweet.content])
```

Gambar 3.4 Proses Pengambilan Data

Fungsi yang diberi nama *tweets_list2* bertujuan untuk mengambil data *tweet*. Dalam pengambilan data menggunakan *snsrape* yang akan disimpan kedalam format *csv* memiliki beberapa bentuk variasi seperti URL, angka, *hashtag*, *datetime*, *username* yang akan dilakukan pengolahan data agar didapatkan hasil *tweet* yang sesuai. Berikut merupakan contoh data *tweet* yang telah didapatkan dapat dilihat dalam Tabel 3.1.

Tabel 3.1 Contoh Data Tweet

No.	Datetime	Text
1.	2022-03-13 12:51:12+00:00	@Wadas_Melawan Jangan sampai Di Tindas. #WadasMelawan#CabutPLWadas
2.	2021-04-23 15:25:52+00:00	Ujian kesabaran yang gimana lagi yang harus kuterima ? ? Beginikah sikap atas jawabanmu yang katanya mengayomi masyarakat ? #WadasMelawan#SaveWadas#CabutPLWadas https://t.co/pit632azzu
3.	2022-06-18 04:38:32+00:00	@RadenKianSayang @PialaSoe @blankbanget @didik_bung @patriot016610 @aniesbaswedan Astagfirullah #SAVEWADAS #WadasMelawan#WadasTolakTambang

4.	2022-06-17 18:36:11+00:00	Ganjar Akan Sulit Mengatasnamakan Rakyat, Sebab Dia Sendiri Sudah Tak Peduli Terhadap Rakyatnya Sendiri Yang Disambut Dengan Kawat. #SAVEWADAS https://t.co/AekJoXeMDI
5.	2022-06-15 09:45:29+00:00	Keren, 4 Mahasiswa UNS Bentangkan Poster #SaveWadas di Depan Ganjar https://t.co/zCRadQKm2W https://t.co/oW0vOv9OGa
6.	2022-06-01 14:12:12+00:00	@GunRomli #Savewadas, elu diem jateng banjir romlah
7.	2022-05-28 17:16:01+00:00	@ganjarpranowo waduh lah yg bikin kebijakan gmna loh masa cuma bisa treak treak tpi gk bisa berbuat apa2, sdh gk jamannya pecitraan. kinerjamu gimana @ganjarpranowo #SaveWadas
8.	2022-05-23 14:19:24+00:00	@CNNIndonesia @msaid_didu Doa dr org2 wadas #savewadas
9.	2022-05-19 17:44:10+00:00	ApakabarNusantara#wdasamelawan#savewadas #wadastolaktambang https://t.co/L6BZ1aM9HX
10.	2022-05-06 06:06:58+00:00	Enggak pilih lah #SaveWadas https://t.co/TfsUhxWCni

Pada Tabel 3.1 menunjukkan hasil dari pengambilan data *tweet* yang memiliki komponen tidak perlu dipergunakan dalam analisis sentimen. Terdapat data *tweet* yang masih memiliki karakter tidak perlu untuk digunakan, sehingga perlu dilakukan data *preprocessing* sehingga didapatkan hasil yang lebih baik.

3.3.2 Data Preprocessing

Data *Preprocessing* merupakan teknik data mining untuk mengubah data mentah menjadi format yang dapat dimengerti. Dalam data *Preprocessing* diperlukan *library python* untuk mempermudah menyelesaikan tugas. *Numpy* merupakan *library Preprocessing* yang dipergunakan untuk perhitungan numrik serta mendukung komputasi, untuk pembersihan atau *cleaning* menggunakan *library pandas*, yang berfungsi dalam proses pengolahan bahas *natural* dalam sebuah teks menggunakan *nlTK*, untuk melakukan proses penyeragaman dalam teks menggunakan *nlTK.corpus*, *nlTK.tokenize* sendiri dipergunakan untuk proses pendukung *tokenizing* dan *stemming*. Sebelum melakukan data *Preprocessing* dapat mengimport beberapa *library* seperti dalam Gambar 3.5

```
import pandas as pd
import numpy as np
import nltk
import string
import re
import emoji
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
```

Gambar 3.5 Library Preprocessing Data

Dalam melakukan data *preprocessing* terdapat beberapa sbu proses yang terdiri dari :

1. Data *cleansing*.

Pada tahap data *cleansing* ini berfungsi untuk menghapus dan membersihkan karakter asing seperti hapus (@), hapus hashtag (#), hapus URL, hapus tanda baca, hapus angka dan hapus *emoticon*. Kode yang digunakan dalam melakukan data cleansing dapat di Lihat dalam Gambar 3.6 dan Gambar 3.7.

```
def remove(tweet):
    tweet = re.sub(r'^RT[\s]+', '', tweet)
    tweet = re.sub(r'#', '', tweet)
    tweet = re.sub(r'http\S+', '', tweet)
    tweet = re.sub(r'[0-9]+', '', tweet)
    tweet = re.sub(r'\_', " ", tweet)
    tweet = re.sub(r"\W+", " ", tweet)
    tweet = re.sub(r'\n', " ", tweet)
    tweet = emoji.get_emoji_regexp().sub("", tweet).strip()
    tweet = re.sub(r':', '', tweet)
    return tweet
```

Gambar 3.6 Tahapan Data Cleansing

Fungsi *remove* diatas dipergunakan dalam penghapusan beberapa karakter yang dianggap tidak diperlukan seperti RT, *hashtag*, URL, angka, *underscore*, tanda baca, *slang* dan *emoticon*.


```
df['clean_tweet'] = np.vectorize(remove_pattern)(df['Text'], "@[\w]*")
df
```

Gambar 3.7 Tahapan Penghapusan Add

Fungsi diatas diberi nama *remove_pattern* yang bertujuan untuk menghapus karakter @ yang tidak diperlukan.

2. Tokenizing.

Tokenizing dipergunakan dalam pemisahan kata dengan menggunakan *whitespace*. *Tokenizing* memiliki berbagai istilah *term* sehingga sering diartikan kata. Kode yang digunakan dalam melakukan *tokenizing* dapat dilihat dalam Gambar 3.8.

```
def tokenizing(tweet):
    tokens = word_tokenize(str.remove(tweet))
    return ' ', join(tokens)
```

Gambar 3.8 Tahapan Tokenizing

Pada fungsi *tokenizing* tersebut berfungsi untuk memisahkan sebuah kata dengan menggunakan *whitespace*.

3. Case Folding.

Case folding bertujuan untuk mengubah penggunaan huruf kapital (*uppercase*) menjadi huruf kecil (*lowercase*). Berikut merupakan kode yang dipergunakan dalam melakukan *case Folding* dapat dilihat dalam Gambar 3.9.

```
def uppercase():
    data = df['clean_bersih'].str.lower()
    return data
```

Gambar 3.9 Tahapan Case Folding

Fungsi *uppercase* diatas bertujuan mengubah sebuah kata yang memiliki huruf kapital (*uppercase*) menjadi huruf kecil (*lowercase*) atau menjadi kecil semua.

4. Stopword

Stopword merupakan proses penghapusan sebuah kata yang tidak digunakan atau tidak penting dan memiliki informasi rendah disebuah teks (“dan”, ”jikalau”, “pun”, “di”, “yang” dan sebagainya). Berikut merupakan ode yang dipergunakan dalam melakukan *Stopword* dapat dilihat dalam Gambar 3.10.

```
from Sastrawi.StopwordRemover.StopwordRemoverFactory import StopwordRemoverFactory

factory = StopwordRemoverFactory()
more_stopword = ['anjing', 'anjir', 'pantek', 'bokep', 'bangsat', 'bok', 'p', 'pulah', 'ehhh', 'sok', 'ban', 'to', 'esdorem', 'wan', 'elso',
'dakik', 'hahahaha', 'wasweswos', 'amp', 'ajig', 'bajingan', 'jembad', 'bajinguk', 'bacot', 'mencret', 'hmmm', 'njirr',
'hahahahaha', 'goblokkk', 'taik', 'bgd', 'dancuk', 'bangsaaadd', 'bacodnya', 'fvck', 'taek', 'asu', 'matamu', 'brengek',
'anjay', 'peleerrr', 'bgcd', 'jir', 'fuck', 'eek', 'cuk', 'genjik', 'prix', 'uhuyyy', 'haq', 'anjing2', 'ckckckkkk',
'hmmm', 'ttt', 'fkg', 'm3', 'kss', 'weh', 'cup', 'uwow', 'wkwekekekekesemka', 'oiii', 'bor', 'imm', 'hehe', 'dex', 'haha',
'wkwkwk1', 'hoho', 'hooohoooo', 'owh', 'wkwkw', 'wkwkwkwkwk', 'upsssttt', 'ckck', 'wii', 'jancoooooook', 'cot', 'ohh']

stopword = factory.create_stop_word_remover()
stopwords = factory.get_stop_words()+more_stopword
print(stopwords)
```

Gambar 3.10 Tahapan Stopword

Dalam *stopword* terdapat kata-kata yang dihasilkan kemudian akan dihapus sesuai dengan *library sastrawi* serta akan ada penambahan beberapa kata. Hasil daftar kata yang akan dipergunakan dapat dilihat dalam Tabel 3.2

Tabel 3.2 Daftar Kata yang Ditambahkan

Daftar Kata
['kita', 'hmmm', 'njirr', 'hahahahaha', 'goblokkk', 'taik', 'bgd', 'dancuk', 'bangsaaadd', 'bor', 'imm', 'hehe', 'dex', 'haha', 'wkwkwk1', 'hoho', 'hooohoooo', 'owh', 'wkwkw', 'wkwkwkwkwk', 'upsssttt', 'ckck"dengan', 'akan', 'juga', 'ada', 'mereka', 'sudah', 'saya', 'terhadap', 'secara', 'agar', 'lain', 'anda', 'begitu', 'mengapa', 'bacodnya', 'fvck', 'taek', 'asu', 'matamu', 'brengek', 'anjay', 'peleerrr', 'bgcd', 'jir', 'fuck', 'eek', 'cuk', 'genjik', 'prix', 'uhuyyy', 'haq', 'anjing2', 'ckckckkkk', 'hmmm', 'ttt', 'fkg', 'm3', 'kss', 'weh', 'cup', 'uwow', 'wkwekekekekesemka', 'oiii', 'kenapa', 'yaitu', 'yakni', 'daripada', 'itulah', 'lagi', 'maka', 'tentang', 'demi', 'dimana', 'kemana', 'pula', 'sambil', 'sebelum', 'sesudah', 'supaya', 'guna', 'kah', 'pun', 'sampai', 'sedangkan', 'selagi', 'sementara', 'tetapi', 'apakah', 'kecuali', 'sebab', 'selain', 'seolah', 'seraya', 'seterusnya', 'tanpa', 'agak', 'boleh', 'dapat', 'dsb', 'dst', 'dll', 'dahulu', 'dulunya', 'anu', 'demikian', 'tapi', 'ingin', 'juga', 'nggak', 'mari', 'nanti', 'melainkan', 'oh', 'ok', 'seharusnya', 'sebetulnya', 'setiap', 'setidaknya', 'sesuatu', 'pasti', 'saja', 'toh', 'ya', 'walau', 'tolong', 'tentu', 'amat', 'apalagi', 'bagaimanapun', 'anjing', 'anjir', 'pantek', 'bokep', 'bangsat', 'bok', 'p', 'pulah', 'ehhh', 'sok', 'bar', 'to', 'esdorem', 'wan', 'elso', 'dakik', 'hahahaha', 'wasweswos', 'amp', 'ajig', 'bajingan', 'jembad', 'bajinguk', 'bacot', 'mencret', 'wii', 'jancoooooook', 'cot', 'ohh']

Setelah muncul kata dalam *stopword*, kemudian akan dilakukan penghapusan kata-kata yang sesuai dengan *trem* yang sesuai dengan daftar kata yang telah dilakukan pada *stopword*. Berikut merupakan kode yang dipergunakan dalam menghapus kata yang sesuai dengan daftar kata dapat dilihat dalam Gambar 3.11.

```
def menghapusKata(tweet):
    clean_list = [word for word in tweet.split() if word not in stopwords]
    return clean_list
```

Gambar 3.11 Tahap Penghapusan Kata

Pada fungsi menghapus kata dilakukan pengecekan kata apakah sudah sesuai dalam *trem* yang sudah ada dalam proses *stopword*. Apabila telah sesuai dengan daftar kata maka dilakukan penghapusan yang telah ditandai sesuai dengan variable *clean_list*.

5. Stemming.

Stemming merupakan suatu proses menemukan kata dasar dalam suatu kata. Kata yang terdapat beberapa imbuhan seperti ber-, per-, di-, ke-, mem-, men-, -an, -kan, dan sebagainya kemudian akan dilakukan proses penghapusan. Kode yang dipergunakan untuk melakukan proses *stemming* dapat dilihat dalam Gambar 3.12.

```
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

factory = StemmerFactory()
stem_say = factory.create_stemmer()

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

term_base = {}

for document in tweet_hapus:
    for term in document:
        if term not in term_base:
            term_base[term] = ""

print(len(term_base))
print("-----")

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

print(term_base)
print("-----")

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

stem_done = tweet_hapus.apply(get_stemmed_term)

print(stem_done)
```

Gambar 3.12 Tahapan Stemming

Proses *stemming* merupakan proses perubahan bentuk *dataframe* kedalam bentuk *dictionary* kemudian akan dilakukan pengubah kata yang ada pada *term* sehingga akan diubah kedalam kata dasar dengan *library StemmerFactory*.

6. Normalization

Normalization merupakan membenaran kata disetiap *term* atau kata penulisannya tidak sesuai dan penggunaan bahasa yang tidak sesuai. Kode yang dipergunakan untuk melakukan *normalization* dapat dilihat dalam Gambar 3.13

```

normalizad_word = pd.read_csv("kata_normalisai.csv")

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)

```

Gambar 3.13 Tahapan Normalization Kata

Normalization merupakan proses membenaran kata tidak tepat yang sesuai dengan daftar kata yang sudah dilakukan pada file *kata_normalisasi.csv* kemudian akan diubah menjadi *dataframe* yang diberi nama *normal_tweet* kemudian akan dihasilkan data *tweet* yang telah *dipreprocessing*. Berikut contoh kata yang digunakan untuk membenaran kata biasa dilihat dalam Tabel 3.3.

Tabel 3.3 Data Normalization

Bifore	After
manoreh	menoreh
adlh	adalah
laah	lah
annama	nama
bu	ibu
emang	memang
opo	apa
kau	kamu
tak	tidak
tau	tahu
yg	yang

Pada tahap *Preprocessing* data akan dihasilkan sebuah data *tweet* yang dapat dipergunakan dan komponennya lebih tertata sehingga akan dilakukan kedalam tahap berikutnya. Berikut merupakan contoh hasil dari *tweet* data yang sudah dilakukan tahap *preprocessing* dapat dilihat dalam Tabel 3.4.

Tabel 3.4 Contoh Data Hasil Preprocessing

No	Datetime	Text
1.	2022-03-13 12:51:12+00:00	jangan tinas wadasmelawan cabutplwadas
2.	2021-04-23 15:25:52+00:00	uji sabar bagaimana terima begini sikap atas jawab kata ayom masyarakat wadasmelawan savewadas cabutplwadas
3.	2022-06-26 02:57:48+00:00	savewadas nyampek jerman dong
4.	2022-06-15 20:33:56+00:00	keren mahasiswa uns bentangkan poster savewadas depan ganjar
5.	2022-02-24 17:12:01+00:00	undang seluruh warga gruduk mabes polri hari jam wib tarik seluruh aparat polri desa wadas stop segala bentuk intimidasi keras warga wadas savewadas wadasmelawan mositidakpercaya
6.	2022-02-10 04:45:30+00:00	salah satu alas tolak warga wadas ialah khawatir ada dampak ekologis mati sumber mata air sangat logis harus dengar sekaligus jamin perintah kamibersamawargawadas savewadas
7.	2022-02-09 05:13:09+00:00	biomu nang instagram mung apus tok pak savewadas wadasmelawan
8.	2022-02-09 23:50:01+00:00	tetap semangat wadasmelawan
9.	2022-02-09 12:06:53+00:00	dulu kamu bela sekarang lewat aparat hak acak percaya inalillahi waina ilaihi rojiun wadasmelawan wadasmelawan
10.	2022-02-08 09:11:08+00:00	doa baik buat saudara wadasmelawan wadasmemanggil wadastolaktambang wadasmelawan

Tabel 3.4 menunjukkan hasil dari *preprocessing* data yang mengubah data *tweet* kedalam bentuk yang lebih baik dan terstruktur sehingga dapat dilakukan kedalam proses berikutnya.

3.3.3 Pelabelan Manual

Pelabelan manual merupakan proses pemberian label kata secara manual di dalam suatu dokumen sehingga dapat dianalisis lebih lanjut mengenai sifat negatif dan positifnya. Berdasarkan data *tweet* yang sudah dilakukan *preprocessing* dan dilakukan pelabelan secara manual sehingga akan mendapatkan data *training* yang berjumlah 888 *tweet* yang terdiri dari 452 data negatif dan 436 data positif. Berikut merupakan hasil pelabelan manual yang telah dilakukan dapat dilihat dalam Gambar 3.14.

	Date	Text	Compound_Score	Sentimen
0	2021-07-15 05:13:59+00:00	warga wadas hasil usir oknum patok ukur illega...	-1	negatif
1	2021-07-15 04:09:48+00:00	oknum laku aksi ukur patok ilegal alas milik w...	-1	negatif
2	2021-07-11 12:01:44+00:00	wadasmelawan savewadas wadastetaplestari kawal...	-1	negatif
3	2021-07-09 04:01:43+00:00	lengkap simak infografis atas narahubung taj m...	-1	negatif
4	2021-07-03 14:25:08+00:00	repost from rilis kaji wadas tipu muslihat bal...	-1	negatif
...
883	2022-02-09 06:12:26+00:00	komnas hak asasi manusia ri merespon ricuh des...	1	positif
884	2022-02-09 06:12:15+00:00	hero m wadasmelawan bongkar semua pimpin citra...	1	positif
885	2022-02-09 06:12:05+00:00	wadasmelawan juang dapat hak rakyat	1	positif
886	2022-02-09 06:10:58+00:00	apa bangun kalau cara rampas hak rakyat sendir...	1	positif
887	2022-02-09 06:08:12+00:00	kemarin suka share berita enggak semoga share ...	1	positif

888 rows x 4 columns

Gambar 3.14 Hasil Pelabelan Manual

Contoh *tweet* data yang sudah dilakukan pelabelan secara manual dapat dilihat dalam Tabel 3.5.

Tabel 3.5 Contoh Tweet Pelabelan Manual

Datetime	Text	Score	Sentimen
2021-07-15 04:09:48+00:00	oknum laku aksi ukur patok ilegal alas milik warga wadas savewadas wadasmelawan usirpematokanilegal usirpengukuranilegal wadastolaktambang usirpengukurilegal	-1	Negatif
2021-05-31 05:42:00+00:00	stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi lewat	-1	Negatif
2021-05-29 02:42:55+00:00	pak bupati stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi lewat	-1	Negatif
2021-05-01 12:44:54+00:00	isi hati warga masyarakat wadas yang bentrok tahan tanah leluhur stop jangan putar balik trus berita savewadas cabutiplwadas wadasmelawan	-1	Negatif
2021-04-24 08:10:01+00:00	semoga cepet sadar allah dengar doa orang didzalimi savewadas	-1	Negatif
2022-03-25 05:23:55+00:00	super bangga sama temen kastradku hebat wadasmelawan untuk bapak ganjar hormat sila baca surat cinta	1	Positif
2022-03-23 13:29:27+00:00	mohon pikirkan dampak kalian ikan buat soal sudut warga usaha tahan hak terus terus intimidasi tuduh sudut didzalimi bagai cara wadasmelawan propagandatur	1	Positif
2022-02-13 14:19:30+00:00	wadasmelawan bagaimana kalau kata enggak keras komnas hak asasi manusia temu fakta keras polisi desa wadas	1	Positif
2022-02-12 08:38:53+00:00	njenengan kok sek pintermen ngeles seng ahli seng hukum legal wadastolaktambang wadasmelawan dosen ugm skema bebas lahan tambang wadas masalah lewat	1	Positif
2022-02-11 00:11:55+00:00	rakyat makin cerdas militan cerah sadar atas hak hak politik lalu dokumenter jaga desa wadas lihat dengar suara warga ceritapingirjalan isupublik wadasmelawan	1	Positif

Dari Tabel 3.5 dapat dilihat *tweet* dengan sentimen positif diberi label nilai 1 sedangkan *tweet* dengan *sentiment* negatif diberi label -1. Pelabelan ini dilakukan untuk memberikan nilai yang bersifat positif dan negatif kemudian dilakukan penghitungan akurasi.

3.3.4 Training Data

Training data merupakan suatu proses *training* data dengan menggunakan metode *Support Vector Machine*. Pada tahap *training* data diawali dengan ekstraksi teks pada data menggunakan proses TF-IDF, selanjutnya akan dilakukan pemodelan klasifikasi menentukan sentimen secara otomatis. Contoh dokumen kata yang akan dilakukan penghitungan TF-IDF manual dapat dilihat dalam Tabel 3.6.

Tabel 3.6 Dokumen TF-IDF

dokumen	Text
d1	pak bupati stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi via
d2	rekam abadi aksi damai wadas unggah mention serta tagar sayabersamawadas savewadas sebar cinta stop intimidasi keras
d3	tahan bentuk cinta paling liar jangan provokasi teman teman savewadas
d4	allah cuma kirim doa semua fatiha savewadas
d5	jaga alam alam jaga hukum alam pak bupati stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi lewat

Tabel 3.6 merupakan contoh kata yang akan dilakukan pembobotan TF-IDF secara manual dengan 5 dokumen yang meliputi d1, d2, d3, d4 dan d5. Dalam pembobotan TF-IDF secara manual ini akan dipergunakan beberapa komponen T yaitu *term* kemudian d yang berarti dokumen d1, d2, d3, d4 dan d5 serta df yaitu banyaknya sebuah *term* yang sering muncul pada suatu dokumen. Contoh penghitungan TF secara manual dapat dilihat dalam Tabel 3.7.

Tabel 3.7 Penghitungan TF Secara Manual.

T	d1	d2	d3	d4	d5	df
pak	1				1	2
bupati	1				1	2
stop	1				1	2
rencana	1				1	2
tambang	1				1	2
batu	1				1	2
andesit	1				1	2
desa	1				1	2
wadas	1	1			1	3
untuk	1				1	2
bangun	1				1	2
bendung	1				1	2
bener	1				1	2
tandatangani	1				1	2
petisi	1				1	2
via	1					1
rekam		1				1
abadi		1				1
aksi		1				1
damai		1				1
unggah		1				1
mention		1				1
serta		1				1
tagar		1				1
sebar		1				1
cinta		1	1			2
keras		1				1
sawewadas	1	1	1	1	1	5
tahan			1			1
bentuk			1			1
paling			1			1
liar			1			1
jangan			1			1
teman			2			1
allah				1		1
cuma				1		1
kirim				1		1
doa				1		1

semua				1		1
fatiha				1		1
alam					3	1
jaga					2	1
hukum					1	1
lewat					1	1
Provokasi		1			1	
intimidasi	1				1	
sayabersamawadas	1				1	

Tabel 3.7 menjelaskan distribusi kemunculan *term* disetiap dokumen. Penghitungan IDF merupakan beberapa komponen *term* yang muncul atau d serta df merupakan jumlah kemunculan *term* dari semua dokumen. Contoh penghitungan IDF dapat dilihat dalam Tabel 3.8.

Tabel 3.8 Penghitungan IDF

T	Df	d/df	Idf
pak	2	2.5	0.39794
bupati	2	2.5	0.39794
stop	2	2.5	0.39794
rencana	2	2.5	0.39794
tambang	2	2.5	0.39794
batu	2	2.5	0.39794
andesit	2	2.5	0.39794
desa	2	2.5	0.39794
wadas	3	1.666667	0.221849
untuk	2	2.5	0.39794
bangun	2	2.5	0.39794
bendung	2	2.5	0.39794
bener	2	2.5	0.39794
tandatangan	2	2.5	0.39794
petisi	2	2.5	0.39794
via	1	5	0.69897
rekam	1	5	0.69897
abadi	1	5	0.69897
aksi	1	5	0.69897
damai	1	5	0.69897
unggah	1	5	0.69897
mention	1	5	0.69897

serta	1	5	0.69897
tagar	1	5	0.69897
sayabersamawadas	1	5	0.69897
sebar	1	5	0.69897
cinta	2	2.5	0.39794
intimidasi	1	5	0.69897
keras	1	5	0.69897
savewadas	5	1	0
tahan	1	5	0.69897
bentuk	1	5	0.69897
paling	1	5	0.69897
liar	1	5	0.69897
jangan	1	5	0.69897
provokasi	1	5	0.69897
teman	1	5	0.69897
allah	1	5	0.69897
cuma	1	5	0.69897
kirim	1	5	0.69897
doa	1	5	0.69897
semua	1	5	0.69897
fatih	1	5	0.69897
alam	1	5	0.69897
jaga	1	5	0.69897
hukum	1	5	0.69897
lewat	1	5	0.69897

Tabel 3.8 merupakan penjabaran hasil dari perhitungan IDF yang dilakukan manual dengan menggunakan persamaan (1). Selanjutnya penghitungan TF-IDF dapat dilihat dalam Tabel 3.9.

Tabel 3.9 Penghitungan TF-IDF

T	d1	d2	d3	d4	d5
pak	0.39794				0.39794
bupati	0.39794				0.39794
stop	0.39794				0.39794
rencana	0.39794				0.39794
tambang	0.39794				0.39794
batu	0.39794				0.39794
andesit	0.39794				0.39794
desa	0.39794				0.39794

wadas	0.221849	0.221849			0.221849
untuk	0.39794				0.39794
bangun	0.39794				0.39794
bendung	0.39794				0.39794
bener	0.39794				0.39794
tandatangan	0.39794				0.39794
petisi	0.39794				0.39794
via	0.69897				
rekam		0.69897			
abadi		0.69897			
aksi		0.69897			
damai		0.69897			
unggah		0.69897			
mention		0.69897			
serta		0.69897			
tagar		0.69897			
sayabersamawadas		0.69897			
sebar		0.69897			
cinta		0.39794	0.39794		
intimidasi		0.69897			
keras		0.69897			
sawewadas					
tahan			0.69897		
bentuk			0.69897		
paling			0.69897		
liar			0.69897		
jangan			0.69897		
provokasi			0.69897		
teman			1.39794		
allah				0.69897	
cuma				0.69897	
kirim				0.69897	
do				0.69897	
semua				0.69897	
fatih				0.69897	
alam					2.09691
jaga					1.39794
hukum					0.69897
lewat					0.69897

Tabel 3.9 merupakan hasil dari penghitungan TF-IDF secara manual dengan menggunakan persamaan (2).

Pada penelitian ini dilakukan klasifikasi dengan menggunakan fitur ekstraksi pembobotan TF-IDF yang akan dihasilkan perhitungan yang dilakukan secara otomatis pada pembobotan *term* disetiap dokumen data *training*. Dalam penghitungan TF-IDF ini menggunakan salah satu *library* yang terdapat pada *Python* yaitu *Sklearn* dan *TfidfVectorizer* serta didukung dengan *library Support Vector Machine* yang akan mendapatkan hasil secara otomatis. Hasil dari perhitungan TF-IDF yang dilakukan secara otomatis dapat dilihat dalam Gambar 3.15 dan Gambar 3.16.

(0, 3141)	0.3900433180379897
(0, 3139)	0.3900433180379897
(0, 3186)	0.07423586941990126
(0, 2564)	0.14310924266518776
(0, 999)	0.34664006775970013
(0, 74)	0.2275345569247124
(0, 3002)	0.37862676988978433
(0, 281)	0.3212507939374856
(0, 1109)	0.3098342457892803
(0, 2964)	0.2778475436591961
(0, 3179)	0.23256083829329569
(0, 3226)	0.15481771513669298
(1, 1027)	0.35210694319440194
(1, 3103)	0.28102316530776167
(1, 2139)	0.3613182406478988
(1, 2046)	0.3370896110889253
(1, 3137)	0.3613182406478988
(1, 938)	0.29567049528388845
(1, 3141)	0.37221289316383777
(1, 3139)	0.37221289316383777
(1, 3186)	0.07084225380992949
(1, 2564)	0.13656715238410294
(1, 3179)	0.11096478064690751
(1, 3226)	0.14774038420632749
(2, 1026)	0.22619952103468804

Gambar 3.15 Perhitungan TF-IDF Awal

(930, 132)	0.15014690418652485
(930, 2374)	0.23205956863560112
(930, 943)	0.1923299760648313
(930, 237)	0.14018967841314559
(930, 3186)	0.049781644481272244
(931, 2683)	0.441682924834104
(931, 1053)	0.2318733308835746
(931, 2339)	0.23886489422994775
(931, 1682)	0.18204916610487265
(931, 2783)	0.17280122073502038
(931, 2805)	0.17280122073502038
(931, 3189)	0.1521320895329167
(931, 721)	0.17145555876458424
(931, 2016)	0.16535583633882756
(931, 2780)	0.18383308920154298
(931, 876)	0.17871250331646937
(931, 2985)	0.24742187620060463
(931, 1130)	0.2584537446671272
(931, 335)	0.4637466617671492
(931, 2640)	0.15817675461805936
(931, 1316)	0.19673593512936507
(931, 3196)	0.17145555876458424
(931, 3201)	0.11081047459202718
(931, 3186)	0.045462496797152724
(931, 3179)	0.0712108341201547

Gambar 3.16 Perhitungan TF-IDF Akhir

Gambar 3.15 dan Gambar 3.16 menunjukkan hasil perhitungan TF-IDF secara otomatis dengan memanfaatkan *library Sklearn* dan *TfidfVectorizer*. Setelah dilakukan perhitungan TF-IDF secara otomatis kemudian dilanjutkan dengan penyesuaian model *dataset* dengan menggunakan algoritma *Support Vector Machine*. Dalam pembuatan model ini digunakan sebuah fungsi untuk mempermudah saat melakukan pemanggilan dan eksekusi pada proses selanjutnya sehingga akan jauh lebih efektif dan efisien. Dalam pembuatan model ini menggunakan kode yang dapat dilihat dalam Gambar 3.17.

```
# pembuatan model dataset svm
SVM = SVC(C=1.0, kernel='linear', degree=3, gamma='auto')
text_pickle=SVM.fit(Train_X_Tfidf,Train_Y)
```

Gambar 3.17 Pembuatan Model Klasifikasi

Setelah dilakukan pemodelan dilakukan pemberian nama *text_pickle* kemudian disimpan kedalam file *pickle* sehingga dapat dipanggil kembali serta dipergunakan. Kode untuk menyimpan dan memanggil file *pickle* yang sudah dilakukan dapat dilihat dalam Gambar 3.18 dan Gambar 3.19.

```
files = open('SVM_classifier_24.pickle', 'wb')
pickle.dump(text_pickle, files)
files.close()
```

Gambar 3.18 Menyimpan File Pickle

```
model = open('SVM_classifier_24.pickle', 'rb')
svm_classifier = pickle.load(model)
svm_classifier
```

Gambar 3.19 Membuka File Pickle

File *pickle* ini kemudian akan dipergunakan dalam eksekusi data *training* dengan jumlah 888 *tweet* yang diambil sebagai *sample* dengan menggunakan *library smote*. Kode yang dipergunakan untuk mengimport *library smote* dapat dilihat dalam Gambar 3.20.

```
from imblearn.over_sampling import SMOTE
```

Gambar 3.20 Import Library SMOTE

Library smote dipergunakan dalam menduplikasi sebuah data sama yang nantinya akan membuat data *sample* yang sama dengan data asli dari kelas minoritas serta untuk menyeimbangkan dataset. Kemudian akan dilakukan *Cross Validation* untuk menghitung akurasi beberapa kali (*k-fold*) yang menggunakan parameter sama. Pada proses ini sering dipergunakan dalam pencarian sebuah nilai akurasi dengan cara melakukan beberapa uji coba sehingga akan mendapatkan tingkat performa nilai data. Kode yang dipergunakan saat melakukan proses *cross validation* dapat dilihat dalam Gambar 3.21.

```
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))
```

Gambar 3.21 Perhitungan Cross Validation

Pada Gambar 3.21 akan di lakukan perhitungan *accuracy* dan *f1s* untuk mengetahui nilai akurasinya data *training*.

3.3.5 Testing

Testing merupakan tahapan dalam mengukur tingkat akurasi model yang sudah dilakukan, tahapan *testing* dipergunakan untuk memprediksi sebuah label dari data uji yang telah ada. Pada tahap *testing* ini akan dilakukan pelabelan secara manual dengan jumlah 400 data *tweet* dengan rincian sentimen negatif sebanyak 225 dan sentiment dengan positif 175 nantinya pelabelan manual ini akan dibandingkan dengan pelabelan secara otomatis yang telah dilakukan oleh model *support vector machine*. Contoh pelabelan secara manual yang akan digunakan untuk perbandingan bisa dilihat pada Tabel 3.10.

Tabel 3.10 Pelabelan Secara Manual

Datetime	Text	Compound_Score	Sentimen
2022-06-15 06:47:27+00:00	hoax yang mana kon tet savewadas	-1	Negatif
2022-04-26 14:20:20+00:00	elo enggak tahu lapang savewadas bikin dengan kanyut laut nusakambangan chat fake aja elo percaya polri bikin saiapa chat fake tersebut bukti memang orang goblok gampang begoin	-1	Negatif
2022-04-16 11:42:52+00:00	haaa alas enggak savewadas	-1	Negatif
2022-02-25 07:42:45+00:00	judul humor kader turun selera politik isu rakyat isu agama warning politik selimut agama lahir tindak sekte baru agama sebut islam sekte aspek anti speker sekte pro speker savewadas humorpublik	-1	Negatif
2022-02-14 11:42:31+00:00	padahal pak kapolres dongeng dulu pernah kata hoax baik hoax versi buzzerp kuasa asuh kakak bina lagi ikan busuk kan memang mulai kepala faktawadas wadasmelawan	-1	Negatif

	wadastolaktambang wadasoradidol savewadas		
2022-02-09 00:43:07+00:00	sekali hak atas ruang hidup esensial tanah banyak orang bukan dar bidang datar unsur hara yang diperjualbelikan semata nilai sejarah nilai hidup nilai lingkung masa depan terus juang wadasmelawan	1	Positif
2022-02-09 00:34:40+00:00	narasi berita akun kok beda banget meski gimanapun karena polemik wadas sudah lama n belum kelar yang gue tahu banyak masyarakat desa wadas yang intimidasi meski mayoritas dukung saya support wadasmelawan	1	Positif
2022-02-09 00:33:09+00:00	bagaimana rakyat harap tega hak asasi manusia kepada kuasa bila nyata langgar hak asasi manusia nya kuasa sendiri rakyat lebih ingin ada waduk kedzaliman spy banjir racun hidup negara wadas wadasmelawan wadasmelawan	1	Positif
2022-02-09 00:28:01+00:00	instagram ganjar pranowo banjir komentar save wadas warga kabar tangkap wadasmelawan ganjar giring motogp	1	Positif
2022-02-08 23:53:18+00:00	pak bantu rakyat bela hak hak rakyat bantu sama sama jiwa raga engkau mati bela rakyat engkau tetap hidup hati rakyat sejarah catat wadasmelawan wadasmemanggil	1	Positif

Tabel 3.10 menunjukkan *compound_score* -1 bersifat *negatif* sedangkan *compound_score* 1 bersifat positif kemudian data tersebut akan dibandingkan dengan data *tweet* yang telah diberi label secara otomatis dengan menggunakan

model *support vector machine*. Kode untuk pelabelan secara otomatis dapat dilihat dalam Gambar 3.22.

```
for i in range(len(prediction_linear)):
    if(prediction_linear[i]==1):
        sentiment_result='positive'
    elif(prediction_linear[i]==-1):
        sentiment_result='negative'
# result_tweet.append({'class':prediction_linear[i], 'result_svm':sentiment_result})
result_tweet.append({'Date': pd_tweet['Date'][i], 'tweet': pd_tweet['Text'][i], 'class':prediction_linear[i]})
```

Gambar 3.22 Pelabelan Secara Otomatis

Gambar 3.22 merupakan *classifier_linear* merupakan pembuatan model untuk melakukan *testing* data *tweet*. Contoh data *tweet* yang telah diberi label secara *otomatis* dengan menggunakan model *support vector machine* dapat dilihat pada Tabel 3.11.

Tabel 3.11 Data Pelabelan Otomatis

Datetime	Text	Score	Sentimen
2022-06-15 06:47:27+00:00	hoax yang mana kon tet savewadas	-1	Negatif
2022-02-25 07:42:45+00:00	judul humor kader turun selera politik isu rakyat isu agama warning politik selimut agama lahir tindak sekte baru agama sebut islam sekte aspek anti speker sekte pro speker savewadas humorpublik	1	Positif
2022-02-25 06:48:04+00:00	gabung diskusi via zoom meeting wadas tanah surga rampas lost and damage jumat feb pukul wib malam kawan kawan wadasmelawan savewadas wadastolaktambang	-1	Negatif
2022-02-21 12:30:22+00:00	kritisbukansinis wadasmelawan savewadas kamibersamamasyarakatparigimoutong departemen aksi propaganda bidang dinamisasi kampus badan eksekutif mahasiswa fakultas hukum universitas padjajaran kabinet konstelasi adaptif responsif dedikatif	1	Positif
2022-02-20 14:54:43+00:00	pak bupati stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi	-1	Negatif

2022-02-12 11:50:21+00:00	yang satu profesor yang agk ngehang karena gagal jadi cawapres satu lagi jenderal polri yang tengsin karena abis pergok sama warga sama palsu sama layan oligark fu k savewadas savewadas	-1	Negatif
2022-02-10 12:54:22+00:00	die takut makan kalau bela savewadas	1	Positif
2022-02-10 08:14:01+00:00	tahan bentuk cinta paling liar jangan provokasi teman teman savewadas	1	Positif
2022-02-10 06:26:23+00:00	apun alas jangan rampas tanah rakyat savewadas	1	Positif
2022-02-10 04:22:43+00:00	mayak anak buah e nek pingin ngopi po ngeteh kon sangu dewe bang savewadas wadasmelawan	-1	Negatif

Tabel 3.11 menunjukkan hasil dari pelabelan secara otomatis dengan model *Support Vector Machine*. Selanjutnya data *tweet* yang diberi label secara manual dan data *tweet* yang telah diberi label secara *otomatis* digabungkan untuk dilakukan pengecekan tingkat akurasi model *Support Vector Machine*. Contoh data *tweet* yang telah dilakukan penggabungan dapat dilihat dalam Tabel 3.12.

Tabel 3.12 Data Gabungan

Datetime	Text	Actual	Predik
2022-06-15 06:47:27+00:00	hoax yang mana kon tet savewadas	negatif	negatif
2022-02-25 07:42:45+00:00	judul humor kader turun selera politik isu rakyat isu agama warning politik selimut agama lahir tindak sekte baru agama sebut islam sekte aspek anti speker sekte pro speker savewadas humorpublik	negatif	positif
2022-02-25 06:48:04+00:00	gabung diskusi via zoom meeting wadas tanah surga rampas lost and damage jumat feb pukul wib malam kawan kawan wadasmelawan savewadas wadastolaktambang	negatif	negatif
2022-02-21 12:30:22+00:00	kritisbukansinis wadasmelawan savewadas kamibersamamasyarakatparigimoutong	negatif	positif

	departemen aksi propaganda bidang dinamisasi kampus badan eksekutif mahasiswa fakultas hukum universitas padjadjaran kabinet konstelasi adaptif responsif dedikatif		
2022-02-20 14:54:43+00:00	pak bupati stop rencana tambang batu andesit desa wadas untuk bangun bendung bener savewadas tandatangani petisi	negatif	negatif
2022-02-12 11:50:21+00:00	yang satu profesor yang agk ngehang karena gagal jadi cawapres satu lagi jenderal polri yang tengsin karena abis pergok sama warga sama palsu sama layan oligark fu k savewadas savewadas	negatif	negatif
2022-02-12 12:08:40+00:00	malam hari mobil polisi datang desa wadas bentuk teror rezim jokowi warga covid sabtu februari tinggal hello handsome tahun jamin hari tua malem minggu malam minggu wadastolaktambang wadastolaktambang	positif	negatif
2022-02-08 17:18:17+00:00	wadasmelawan wadasmelawan jabat rampas hak rakyat	positif	Positif
2022-02-08 18:29:39+00:00	tindak kalian sengaja undang seluruh rakyat biasa guna hak cari benar sendiri konsolidasi rakyat biasa kepong kota wadasmelawan wadasadalahkita	Positif	negatif
2022-02-08 05:24:35+00:00	hati tahan sekarang wilayah hidup suka rampas lalu apa alas hidup wadas lawan wadasmemanggil wadasoradidol wadasmelawan savewadas	negatif	positif

Pada fungsi ini akan memberi keterangan *True Negatif (TN)*, *False Positif (FP)*, *True Positif (TP)* dan *False Negatif (FN)* disetiap data *tweet*. Kode yang digunakan untuk melakukan pemberian keterangan dapat dilihat dalam Gambar 3.23.

```
df.loc[(df['Sentimen'] == 'positif') & (df['class'] == 'positif'), 'keterangan'] = 'TP'
df.loc[(df['Sentimen'] == 'positif') & (df['class'] == 'negatif'), 'keterangan'] = 'FP'
df.loc[(df['Sentimen'] == 'negatif') & (df['class'] == 'negatif'), 'keterangan'] = 'TN'
df.loc[(df['Sentimen'] == 'negatif') & (df['class'] == 'positif'), 'keterangan'] = 'FN'
```

Gambar 3.23 Menentukan TN, FN, TP dan TN

Pada fungsi ini akan dilakukan pemberian keterangan jika *actual* bersifat *negatif* dan *predik* bersifat *negatif* maka menjadi *True Negatif (TN)*, jika *actual* bersifat *positif* dan *predik* bersifat *negatif* maka menjadi *False Positif (FP)*, jika *actual* bersifat *positif* dan *predik* bersifat *positif* maka menjadi *True Positif (TP)* dan jika *actual* bersifat *negatif* dan *predik* bersifat *positif* maka menjadi *False Negatif (FN)*. Hasil pemberian keterangan dapat dilihat dalam Gambar 3.24.

	Date	Text	Compound_Score	Sentimen	class	keterangan
0	2022-06-15 06:47:27+00:00	hoax yang mana kon tet savewadas	-1	negatif	negatif	TN
1	2022-05-23 14:19:24+00:00	doa dari org wadas savewadas	-1	negatif	negatif	TN
2	2022-04-26 14:20:20+00:00	elo enggak tahu lapang savewadas bikin dengan ...	-1	negatif	negatif	TN
3	2022-04-25 21:04:11+00:00	doa warga wadas doa kok buat akun yucub savewadas	-1	negatif	negatif	TN
4	2022-04-16 11:42:52+00:00	haaa alas enggak savewadas	-1	negatif	negatif	TN
...
395	2022-02-08 09:40:43+00:00	tindak keras yang selalu utama halo wadastolak...	1	positif	negatif	FP
396	2022-02-08 09:14:40+00:00	orang hak hidup hak tahan hidup hidup uud pasa...	1	positif	positif	TP
397	2022-02-08 06:16:57+00:00	negeri covid sama internet setting takedown ha...	1	positif	positif	TP
398	2022-02-08 05:14:06+00:00	panjang umur juang rebut hak ruang hidup bumi ...	1	positif	positif	TP
399	2022-01-15 11:22:10+00:00	bangun pakai intimidasi keras cari untung proy...	1	positif	negatif	FP

Gambar 3.24 Pemberian Keterangan

Gambar 3.24 merupakan hasil data *tweet* yang sudah dilakukan pemberian keterangan terhadap sentimen *actual* dan sentimen *predik* yang hasilnya *TP*, *TN*, *FP* dan *FN*.

3.3.6 Hasil Klasifikasi

Setelah dilakukan kalsifikasi pemodelan dengan menggunakan *support vector machine* yang telah diuji dengan *training* dan *testing* yang mendapatkan akurasi sesuai. Selanjutnya akan dilakukan klasifikasi dengan menggunakan data keseluruhan dengan total data *tweet* sebanyak 13397. Hasil data keseluruhan yang sudah dilakukan klasifikasi dengan menggunakan metode *support vector machine* dapat dilihat dalam Gambar 3.25.

	Date	tweet	class
0	2022-03-13 12:51:12+00:00	jangan tinas wadasmelawan cabutplwadas	positif
1	2021-04-23 15:25:52+00:00	uji sabar bagaimana terima begini sikap atas j...	positif
2	2022-06-30 00:47:24+00:00	layan oligarki tinas rakyat savewadas	positif
3	2022-06-27 09:39:02+00:00	akhir pekan kemarin anjangsana purworejo bawa ...	positif
4	2022-06-26 02:57:48+00:00	savewadas nyampek jerman dong	positif
...
13393	2021-12-09 08:14:22+00:00	re wadas lawan bahkan akses jalan quarry desa ...	negatif
13394	2021-07-06 09:07:47+00:00	jadi tanggal juni kemarin cerita kirim surat g...	positif
13395	2021-06-03 06:12:40+00:00	kotak nasi solidaritas dapur umum buruhgandong...	negatif
13396	2021-04-29 08:03:07+00:00	ikan semangat dukung partisipasi suara kampany...	negatif
13397	2021-04-29 07:52:36+00:00	konferensi pers lebih sedang langsung silah ce...	positif

Gambar 3.25 Hasil Data Keseluruhan.

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YAN
YOGYAKARTA