

**UJIAN AKHIR SEMESTER DATA MINING II**  
**KLASIFIKASI DATA TWEET TENTANG MINYAK GORENG**  
**MENGGUNAKAN NAÏVE BAYES CLASSIFIER**



Penulis :

Muhammad Zulfikar Reyhan (192103008)

**PROGRAM STUDI SISTEM INFORMASI**  
**FAKULTAS TEKNIK DAN TEKNOLOGI INFORMASI**  
**UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA**

**2022**

## **KATA PENGANTAR**

Dengan hormat,

Puji dan syukur penulis panjatkan kepada Allah SWT atas nikmat sehat yang diberikannya penulis dapat menyelesaikan Ujian Akhir Semester pada mata kuliah Data Mining II ini dengan lancar .

Makalah ini dibuat dalam rangka memenuhi salah satu penilaian mata kuliah Data Mining II dan sebagai dosen pengampu yaitu oleh Bapak Muhammad Habibi, S.Kom., M.Cs. di Universitas Jenderal Achmad Yani Yogyakarta pada tahun akademik 2022.

Ujian Akhir Semester ini berisi laporan praktik klasifikasi data twitter tentang Minyak goreng yang data nya sudah disiapkan oleh dosen, program nya dikerjakan menggunakan Bahasa pemrograman Python dan metode nya Naïve Bayes Classifier.

Ucapan terima kasih kepada dosen pengampu yang telah memberikan masukan, ilmu dan juga praktik yang membuat saya mampu mengerjakan ujian akhir semester ini. Penulis menyadari belum sempurna dalam membuat laporan ini, oleh karena itu penulis mengharapkan kritik dan saran dari Dosen pengampu dan teman – teman yang membaca.

Kurang lebih penulis mohon maaf.

Hormat.

Yogyakarta, 8 Juni 2022

Muhammad Zulfikar Reyhan

## DAFTAR ISI

COVER .....	1
KATA PENGANTAR .....	2
DAFTAR ISI.....	3
DAFTAR GAMBAR .....	4
BAB 1 PENDAHULUAN .....	5
A. Latar belakang .....	5
B. Tujuan.....	5
C. Manfaat.....	5
BAB 2 PEMBAHASAN .....	7
A. Analisis sentiment .....	7
B. Text mining .....	7
C. Text Preprocessing.....	7
D. Klasifikasi.....	7
E. Studi kasus : Minyak Goreng .....	8
BAB 3 KESIMPULAN DAN SARAN .....	13
A. KESIMPULAN .....	13
B. SARAN .....	13
DAFTAR PUSTAKA .....	14

## DAFTAR GAMBAR

Gambar 1 Program python text preproceccing data tweet .....	8
Gambar 2 Hasil text preproceccing tweet .....	9
Gambar 3 Pemberian label dan kelas pada data tweet .....	10
Gambar 4. Data latih/Training .....	10
Gambar 5. Hasil Evaluasi.....	11
Gambar 6. Tweet test .....	12
Gambar 7. Hasil klasifikasi.....	12

## **BAB 1 PENDAHULUAN**

### **A. Latar belakang**

Pada masa modern seperti saat ini media social bukan hanya menjadi tempat berkeluh kesah dan mengeskpos masalah pribadi dan masalah yang tidak serius. Namun, sekarang media social bisa menjadi media berita yang digunakan bukan hanya untuk warga biasa bahkan pemimpin dunia.

Belum lama ini dunia di kagetkan dengan penyerangan tantara rusia ke ukraina, walaupun masalah tersebut adalah masalah internal kedua negara, namun semua pemerintah dunia mengkhawatirkan negara mereka masing masing. Termasuk di negara kita Indonesia.

Di dalam negri permasalahan ini bukan hanya menjadi perhatian pemerintah dalam bersikap dan bersiap. Namun masyarakat media social tweeter Indonesia juga sempat tranding dengan pembahasan perang antara dua negara ini.

Melalui makalah ini penulis akan mengklasifikasikan tweet yang Pro dan Kontra tentang masalah peperangan ini. Penulis akan menggunakan naive bayes classifier dalam mengklasifikasi tweet yang di temukan dari spreadsheet Tag. Diharapkan makalah ini bisa menjadi pembelajaran untuk penulis dan orang lain yang membacanya.

### **B. Tujuan**

Tujuan penulisan makalah ini sebagai berikut :

1. Mengetahui metode Naive Bayes Classifier
2. Mengklasifikasikan Tweet antara Pro dan Kontra
3. Mengetahui Langkah Langkah metode Naive Bayes Classifier
4. Sebagai Tugas Akhir matakuliah Data Mining II

### **C. Manfaat**

Manfaat penulisan makalah ini sebagai berikut :

1. Melihat bagaimana Masyarakat indonesia menyikapi Perang tersebut

2. Membedakan tweet yang Pro dan Kontra terhadap perang tersebut menggunakan metode Naive Bayes Classifier.

## **BAB 2 PEMBAHASAN**

### **A. Analisis sentiment**

Analisis Sentimen adalah salah satu cabang sebuah penelitian text mining, yang berkaitan dengan bidang yang lebih luas seperti pengolahan data kegiatan tertentu (Darwis, Siskawati and Abidin, 2021)

Menurut (Samir *et al.*, 2021) Analisis sentimen adalah proses menentukan sentimen dan mengelompokkan polaritas teks dalam dokumen atau kalimat sehingga kategori dapat ditentukan sebagai sentimen positif, negatif, atau netral.

### **B. Text mining**

Menurut (Samir *et al.*, 2021) Teks mining bertujuan menghasilkan informasi dari satu set dokumen. Text Mining mampu menghasilkan informasi melalui pemrosesan, pengelompokan, dan analisis data-data tidak terstruktur dalam jumlah besar.

### **C. Text Preprocessing**

Teks preprocessing adalah pembersihan, dan penyederhanaan teks agar dapat diproses lebih lanjut (Purwiantono and Aditya, 2020). Berikut ini merupakan tahapan dari text preprocessing (Rozi, Hamdana and Iqbal Alfahmi, 2018):

1. Case folding
2. Cleaning
3. Tokenizing
4. Stop word.

### **D. Klasifikasi**

Data dari Twitter tersebut akan dapat diproses menggunakan text mining, kemudian dilanjutkan dengan mengklasifikasikan Tweet ke dalam dua kelas, Pro dan kontra. Klasifikasi dapat memberikan kemudahan bagi semua pengguna untuk melihat pengguna tweeter yang Pro dan Kontra dengan serangan tersebut.(Darwis, Siskawati and Abidin, 2021)

## E. Studi kasus : Minyak goreng

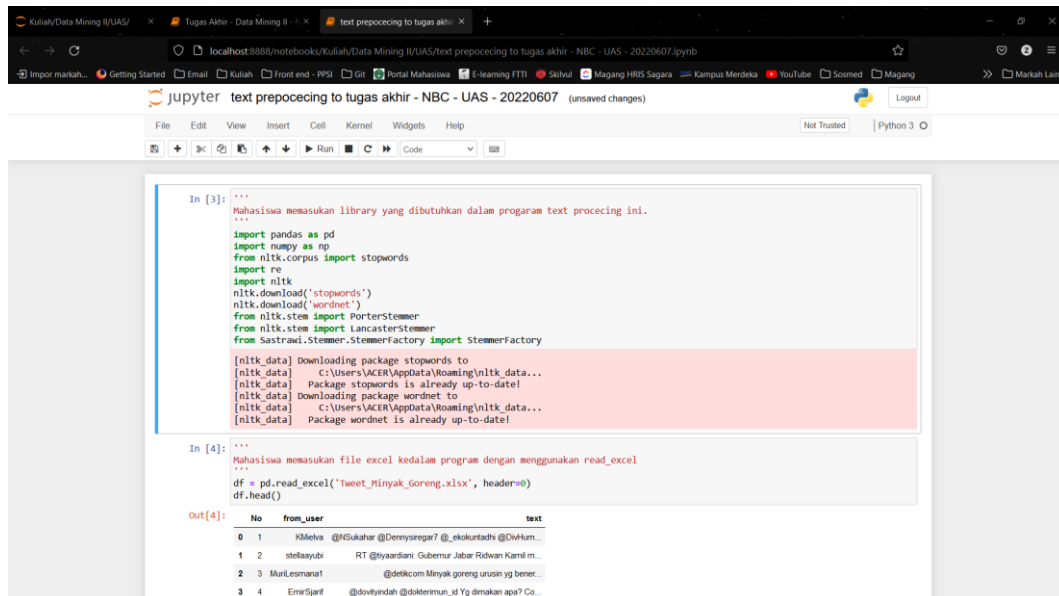
Pada studi kasus ini penulis melakukan proses klasifikasi tweet yang mengangkat berita tentang minyak goreng. Pada saat terjadi kelangkaan dan kenaikan harga minyak goreng yang tidak wajar.

Kasus minyak goreng ini menjadi salah satu kejadian yang bukan hanya membuat ibu rumah tangga menjadi kesulitan, tapi juga membuat sektor UMKM menjadi terkena imbas dari harga minyak goreng yang meningkat dan kelangkaan minyak goreng tersebut.

Proses yang akan dilakukan untuk mengerjakan makalah ini yaitu sebagai berikut:

### 1. Preprocessing data

Pada tahap ini penulis memasukan data dalam bentuk excel kedalam program berbahasa pemograman python. Sebelum tugas akhir ini penulis sudah mengerjakan tugas tugas lain yang sesuai dengan kebutuhan pada tugas akhir ini.



```
In [3]: """
Mahasiswa memasukan library yang dibutuhkan dalam program text prececing ini.
"""
import pandas as pd
import numpy as np
from nltk.corpus import stopwords
import re
import nltk
nltk.download('stopwords')
nltk.download('wordnet')
from nltk.stem import PorterStemmer
from nltk.stem import LancasterStemmer
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\ACER\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\ACER\AppData\Roaming\nltk_data...
[nltk_data] Package wordnet is already up-to-date!

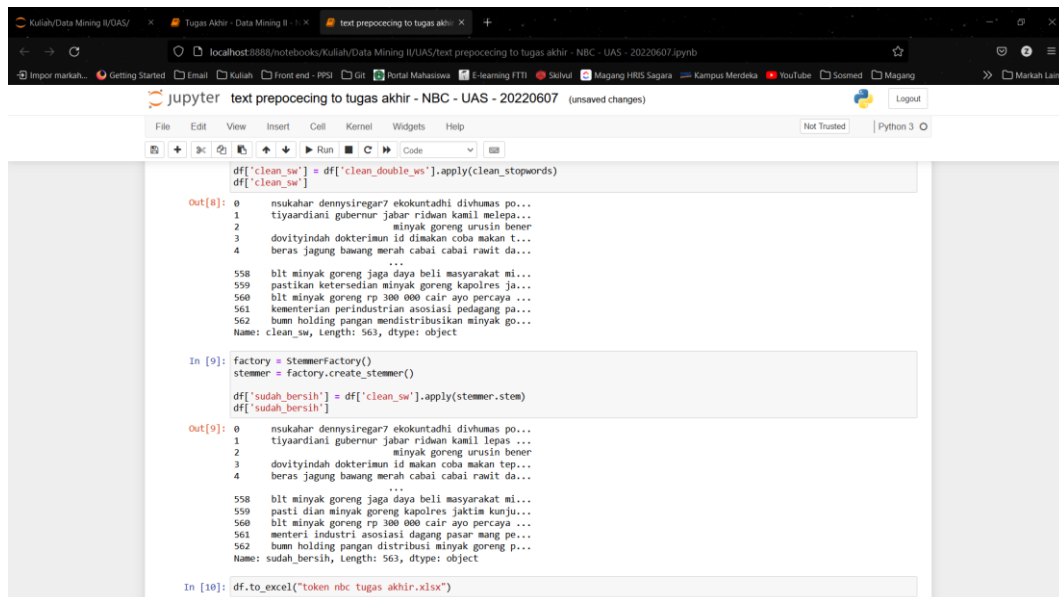
In [4]: """
Mahasiswa memasukan file excel kedalam program dengan menggunakan read_excel
"""
df = pd.read_excel('tweet_minyak_goreng.xlsx', header=0)
df.head()

Out[4]:
```

	No	from_user	text
0	1	Kikiwa @NSukahar @Dennysregar7 @ekokuntadhi @DhHum...	
1	2	stellaayubi RT @nyaaadani: Gubernur Jabar Ridwan Kamil m...	
2	3	MariLesmana1 @dellacom Minyak goreng urusin yg bener...	
3	4	EmreSjani @dovityindah @dokterimun_id Yg dimakan apa? Co...	

Gambar 1 Program python text preproceccing data tweet





```
df['clean_sw'] = df['clean_double_ws'].apply(clean_stopwords)
df['clean_sw']

Out[8]: 0 nsukahar dennysiregar? ekokuntadhi dihumas po...
1 tiyaardiani gubernur jabar ridwan kamil melep...
2 minyak goreng urusin bener
3 dovityindah dokterimun id dimakan coba makan t...
4 beras jagung bawang merah cabai cabai rawit da...
...
558 blt minyak goreng jaga daya beli masyarakat mi...
559 pastikan ketersediaan minyak goreng kapolres ja...
560 blt minyak goreng rp 300 800 cair ayo percaya ...
561 kementerian perindustrian asosiasi pedagang pa...
562 bumn holding pangan mendistribusikan minyak go...
Name: clean_sw, length: 563, dtype: object

In [9]: factory = StemmerFactory()
stemmer = factory.create_stemmer()
df['sudah_bersih'] = df['clean_sw'].apply(stemmer.stem)
df['sudah_bersih']

Out[9]: 0 nsukahar dennysiregar? ekokuntadhi dihumas po...
1 tiyaardiani gubernur jabar ridwan kamil lepa...
2 minyak goreng urusin bener
3 dovityindah dokterimun id makan coba makan tep...
4 beras jagung bawang merah cabai cabai rawit da...
...
558 blt minyak goreng jaga daya beli masyarakat mi...
559 pasti dian minyak goreng kapolres jaktim kunju...
560 blt minyak goreng rp 300 800 cair ayo percaya ...
561 menteri industri asosiasi dagang pasar mang pe...
562 bumn holding pangan distribusi minyak goreng p...
Name: sudah_bersih, length: 563, dtype: object

In [10]: df.to_excel("token nbc tugas akhir.xlsx")
```

*Gambar 2 Hasil text preprocecing tweet*

Pertama data di ubah menjadi huruf kecil, lalu data di bersihkan dari angka symbol dan space kosong, setelah itu kata kata yang tidak ada makna dihapus menggunakan stopwords, setelah itu baru data dikeluarkan lagi dengan format excel.

## 2. Pelabelan tweet

Pada tahap setelah data dilakukan preprocessing, data kemudian dilabel berdasarkan konten yang terkandung didalamnya, jika konten itu berunsur baik, solusi, kabar baik, dan tidak menghina maka akan dilabeli Positif, namun jika sebaliknya maka akan dilabeli Negatif.

Saat proses pelabelan ditemukan satu issue dimana ada data tweet yang sama dengan pembuat yang berbeda. Oleh sebab itu penulis melakukan pembersihan lagi data yang duplikasi dengan tools dari excel.

	from_user	text	kelas	label
0	KMieha	nsukahar dennysiregar? elokuntadhi divhumas poli ccicpolri lebay sengsara tukang goreng laris manis goreng minyak goreng	NEGATIF	-1
1	MuriLesmana1	minyak goreng urusan bener	NEGATIF	-1
2	akufefel	beras jagung bawang merah cabai cabai rawit daging ayam telur ayam minyak goreng stokpanganaman coa0vbt1bbyy pengamatekonom minyak goreng posisi mei kira	NEGATIF	-1
3	AnaSirehar	kang bojes sebab harga minyak goreng mahal co0nAkqjnmj	NEGATIF	-1
4	rajaInfo1	nicho silalahi paham minyak goreng mahal komisaris ptn model budimandjatmiko ngerjain ngebacot mutu wajar bumn rugi komisaris model co0vgtEyximz	NEGATIF	-1
5	Aa49852008	minyak goreng lanka orang kudet deh stok minyak goreng pasti udah limbah indonesia dukungmendaufuti cowb5ddq8y90	NEGATIF	-1
6	Oktarani5	albumi the nangorian cerita au idup lo udah kayak minyak goreng ayam hipotesa kisah kisah kisah kisah wilayah bandung sumedang co0vbw6f6dd	NEGATIF	-1
7	sumilohilo	glat bhabinkamtibmas kec pematang jaya polsek pangkal susu brigadir rico surya ardana sh kece stok minyak goreng jenis curah kemas kemas premium masyarakat	NEGATIF	-1
8	PkHumas			

Gambar 3 Pemberian label dan kelas pada data tweet

### 3. Pelatihan/Training

Dari 563 data yang diberikan, hanya 216 data tweet yang digunakan untuk menjadi data training. Karena banyak tweet yang ganda dan banyak tweet yang tidak mengandung topik yang sedang dipecahkan oleh penulis. Sisah data digunakan untuk data testing nanti nya.

```

In [2]: df = pd.read_excel(r"data_training_uas.xlsx")
df = pd.DataFrame(df)
df.head()

Out[2]:
   Unnamed: 0  from_user  text  kelas  label
0           0  KMieha  nsukahar dennysiregar? elokuntadhi divhumas po...  NEGATIF  -1
1           1  MuriLesmana1  minyak goreng urusan bener  NEGATIF  -1
2           2  akufefel  beras jagung bawang merah cabai cabai rawit da...  NEGATIF  -1
3           3  AnaSirehar  pengamatekonom minyak goreng posisi mei kira  NEGATIF  -1
4           4  rajaInfo1  kang bojes sebab harga minyak goreng mahal cox...  NEGATIF  -1

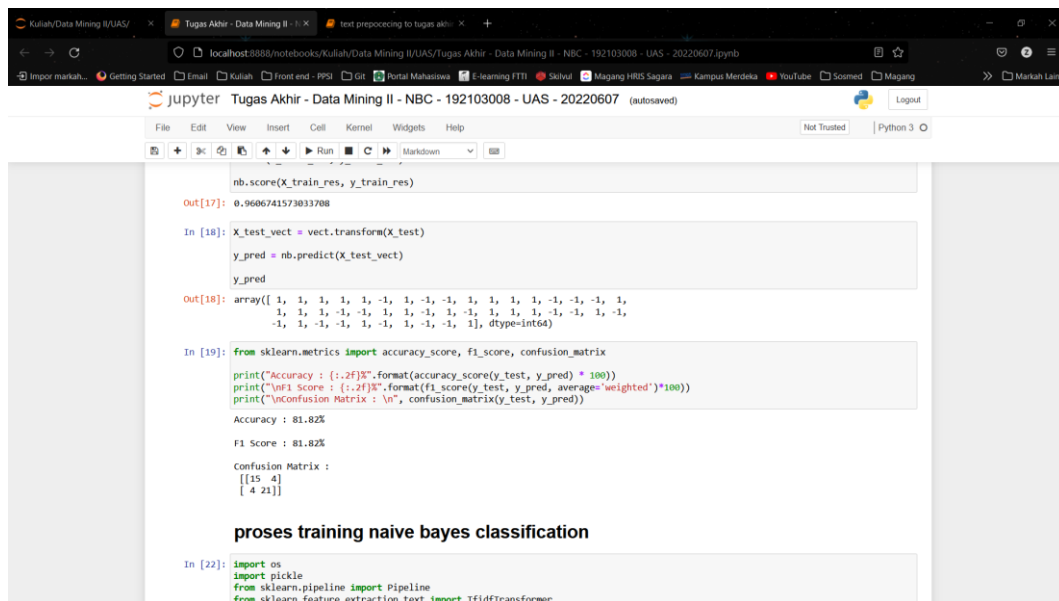
In [3]: df

Out[3]:
   Unnamed: 0  from_user  text  kelas  label
0           0  KMieha  nsukahar dennysiregar? elokuntadhi divhumas po...  NEGATIF  -1
1           1  MuriLesmana1  minyak goreng urusan bener  NEGATIF  -1
2           2  akufefel  beras jagung bawang merah cabai cabai rawit da...  NEGATIF  -1
3           3  AnaSirehar  pengamatekonom minyak goreng posisi mei kira  NEGATIF  -1
4           4  rajaInfo1  kang bojes sebab harga minyak goreng mahal cox...  NEGATIF  -1
...         ...
211          189  EkaFit36468936  warga padle terima sembako uang minyak goreng...  POSITIF  1
212          190  Jmus_Kalmesada  cair sembako uang minyak goreng kantor pos ayo...  POSITIF  1
213          191  Chansul3052684  sukrajati id rhyt rendu era sby listrik murah...  POSITIF  1
214          192  archdamm  fenomena jht car usia pensiun tc 10 juta up b...  POSITIF  1
215          193  alydrusbella1  postifcovid bahagia enak enak vaksin kodm ja...  POSITIF  1
216 rows x 5 columns
  
```

Gambar 4. Data latih/Training

#### 4. Nilai Evaluasi

Hasil pengecekan akurasi pada program ini, penulis mendapatkan hasil akurasi yaitu 81.82%. Hasil tersebut di dapatkan dari modul yang di ambil dari sklearn.metric. Lalu untuk pengecekan F1 Score pada program ini didapatkan hasil yaitu 81.82% dimana dihitung dari presisi dan penarikan kembali test.



```
nb.score(X_train_res, y_train_res)
Out[17]: 0.9606741573033708

In [18]: X_test_vect = vect.transform(X_test)
         y_pred = nb.predict(X_test_vect)
         y_pred

Out[18]: array([[ 1,  1,  1,  1,  1, -1,  1, -1, -1,  1,  1,  1,  1, -1, -1,  1,
                  1,  1,  1, -1,  1,  1, -1,  1,  1,  1,  1, -1, -1,  1, -1,
                  -1,  1, -1,  1, -1,  1, -1, -1,  1], dtype=int64)

In [19]: from sklearn.metrics import accuracy_score, f1_score, confusion_matrix

         print("Accuracy : {:.2F}%".format(accuracy_score(y_test, y_pred) * 100))
         print("F1 Score : {:.2F}%".format(f1_score(y_test, y_pred, average="weighted") * 100))
         print("\nConfusion Matrix : \n", confusion_matrix(y_test, y_pred))

Accuracy : 81.82%

F1 Score : 81.82%

Confusion Matrix :
[[15  4]
 [ 4 21]]

proses training naive bayes classification

In [22]: import os
         import pickle
         from sklearn.pipeline import Pipeline
         from sklearn.feature_extraction.text import TfidfTransformer
```

Gambar 5. Hasil Evaluasi

#### 5. Klasifikasi data testing

Tahap terakhir yaitu penulis memasukan data tweet yang tidak digunakan untuk data training/latih yang akan di proses oleh program untuk mengetahui klasifikasi dari tweet tersebut secara otomatis dengan hasil dari tweet traing sebelumnya.

The screenshot shows a Jupyter Notebook interface with a browser window at the top. The notebook has two cells. The first cell, labeled 'In [72]:', contains a list of tweets assigned to the variable 'tweet\_test'. The second cell, labeled 'In [73]:', contains the code 'pd.DataFrame(tweet\_test)'. The output of the second cell, labeled 'Out[73]:', shows a DataFrame with one row and one column, containing the first tweet from the list.

```
In [72]: tweet_test = ['ayo percaya Jokowi naik harga sembako bmn minyak goreng cousteluJwox',  
                        '86 684 warga pandeglang blt minyak goreng ayo percaya Jokowi codd4sfvry',  
                        'cahyaniika3 blt minyak goreng cair minggu kantor pos ayo percaya Jokowi cofdk4selextf',  
                        'makan bahan minyak goreng alternatif menu sehat pas minyak goreng habis dapur bening tepee bumbu kuning botok teri daun singkong',  
                        'pebak kubu raya unik distribusi minyak goreng curah manfaat data laku umka data sistem informasi geospasial distribusimigor co',  
                        'perintah gelontor rp695 triliun bantu langsung tunai blt minyak goreng ramadan simak skema salur infografis cminidonesia cooyda',  
                        'ayo percaya Jokowi blt minyak goreng cair minggu kantor pos copnqx57y5h4',  
                        'masreghaan sorry zaman habibie beras harga 6000 turun sampe 2000 gusdur minyak goreng nasional gua saudara jual migordt tv harga',  
                        'an asosiasi dagang pasar mang penting presiden Joko Widodo ecek tinjau harga minyak goreng pasar rakyat menteri bumn erick thoiki',  
                        'sobat bumn bumn holding pangan distribusi minyak goreng pasar pasar rakyat 17 juta liter distribusi minyak goreng wujud kolaborasi',  
                        'pasti dian minyak goreng kapolres jaktim kunjung pasar cilubur ciracas cofskuzipre',  
                        'bumn holding pangan distribusi minyak goreng pasar pasar rakyat 17 juta liter distribusi minyak goreng wujud kolaborasi bumn par',  
                        'mantap berkat kerja keras perintah stok minyak goreng curah limbah bestiee dukungmendaglutffi codatknnp1lb',  
                        'giat bhabinkamtibas kel beras basah polsek pangkal susu brigadir j a siregar laksana giat kece sedia minyak goreng kel beras ba',  
                        'henry4ck alingabalinnew alingabalinnew kemendag erickthohir beli minyak goreng subsidi pakai ktp partaisoced cogginoyj16c',  
                        'salah alas dukung ridwankamil beliau solutif beliau luncur aplikasi sapawarga mudah masyarakat beli minyak goreng curah subsidi',  
                        'hujan besti untung salur blt minyak goreng kantor pos antar langsung tugas khawatir dah dukungmendaglutffi cokofsaty03',  
                        'litachania2 alhamdulillah berkah puasa minyak goreng al quran',  
                        'riah acara damai Indonesia tvone kemarin sinar mas riha bantu minyak goreng wakaf alquran serta vaksin booster salut berkah puas',  
                        'realisasi salur blt minyak goreng surabaya capai 72 persen bltminyagoreng ericahyadi kemensos penyaluranbltminyagorengdisuraba',  
                        'mandiraja babinsa koramil 07 mandiraja monitoring salur dana bptn minyak goreng aula balai desa somawangi 1804 kodim banjarnegar',  
                        'ridwan kamil tegus kirin perdana minyak goreng curah pes via aplikasi cornm66ezhji',  
                        'distribusi blt minyak goreng ntb target tuntas lebaran codpi8aie4h',  
                        'daftar blt minyak goreng rp 300 ribu online cosdgc2ome3',  
                        'perintah laku cepat distribusi blt minyak goreng jelang lebaran dukungmendaglutffi colgx3sf38b5',  
                        'lapis masyarakat suka cita blt minyak goreng dukungmendaglutffi coqkyen1uennv',  
                        'renahoston178 rakornas program bantu tunai pangan blt minyak goreng menteri koordinator bidang ekonomi airlangga hrt panglima te',  
                        'salur beras usaha abdul rasyid salur 100 ribu liter minyak goreng seluruh penjuru kalimantan publisherstory cowr3lsetca',  
                        'perintah terbit atur sedia minyak goreng usaha mikro co7tsgeos61i']
```

```
In [73]: pd.DataFrame(tweet_test)
```

```
Out[73]:
```

0	ayo percaya Jokowi naik harga sembako bmn miny...

Gambar 6. Tweet test

The screenshot shows a Jupyter Notebook interface with a browser window at the top. The notebook has two cells. The first cell, labeled 'In [76]:', contains the code 'data=pd.DataFrame(result\_tweet)' and 'data'. The second cell, labeled 'Out[76]:', shows the output of the classification, which is a DataFrame with three columns: 'tweet', 'class', and 'result\_nbc'. The output shows 20 rows of data, with the first row being 'ayo percaya Jokowi naik harga sembako bmn miny...' classified as 'POSITIF'.

```
In [76]: data=pd.DataFrame(result_tweet)
```

```
data
```

```
Out[76]:
```

	tweet	class	result_nbc
0	ayo percaya Jokowi naik harga sembako bmn miny...	1	POSITIF
1	86 684 warga pandeglang blt minyak goreng ayo ...	1	POSITIF
2	cahyaniika3 blt minyak goreng cair minggu kant...	1	POSITIF
3	makan bahan minyak goreng alternatif menu seha...	-1	NEGATIF
4	pebak kubu raya unik distribusi minyak goreng...	1	POSITIF
5	perintah gelontor rp695 triliun bantu langsung...	1	POSITIF
6	ayo percaya Jokowi blt minyak goreng cair ming...	1	POSITIF
7	masreghaan sorry zaman habbie beras harga 600...	1	POSITIF
8	an asosiasi dagang pasar mang penting presiden...	-1	NEGATIF
9	sobat bumn bumn holding pangan distribusi miny...	1	POSITIF
10	pasti dian minyak goreng kapolres jaktim kunju...	1	POSITIF
11	bumn holding pangan distribusi minyak goreng p...	1	POSITIF
12	mantap berkat kerja keras perintah stok minyak...	1	POSITIF
13	giat bhabinkamtibas kel beras basah polsek pa...	1	POSITIF
14	henry4ck alingabalinnew alingabalinnew kemend...	1	POSITIF
15	salah alas dukung ridwankamil beliau solutif b...	1	POSITIF
16	hujan besti untung salur blt minyak goreng kan...	1	POSITIF
17	litachania2 alhamdulillah berkah puasa minyak...	1	POSITIF
18	riah acara damai Indonesia tvone kemarin sinar...	1	POSITIF
19	realisasi salur blt minyak goreng surabaya cap...	1	POSITIF
20	mandiraja babinsa koramil 07 mandiraja monitor...	1	POSITIF

Gambar 7. Hasil klasifikasi

## **BAB 3 KESIMPULAN DAN SARAN**

### **A. KESIMPULAN**

Hasil dari praktikum pengerjaan ini adalah program yang dapat menentukan suatu tweet memiliki stement yang pro atau kontra dengan topik yang dipilih oleh penulis yaitu Perang Rusia dengan Ukraina.

Program yang penulis rancang ini menghasilkan program yang membantu dalam mengelompokkan sebuah tweet termasuk mendukung atau menolak tidakan rusia terhadap ukraina.

### **B. SARAN**

Program ini belum sempurna dan belum memiliki akurasi yang sangat tinggi, program ini juga masih jauh dari kata sempurna dan tentu masih banyak kekurangan baik yang disengaja maupun karena keterbatasan dari penulis sendiri.

Oleh karena itu program ini perlu dikembangkan dan diperbaiki lagi, agar dapat menjadi lebih baik lagi. Maka penulis memiliki beberapa saran yaitu sebagai berikut:

1. Dalam menambang data pada tweeter sudah banyak akun akun yang tidak berkompetibel, oleh karena itu perlu dikurasi lagi.
2. Data yang didapat dengan menggunakan kata kunci bisa juga menarik pula tweet yang membicarakan hal lain, oleh karena itu perlu dikurangi lagi.
3. Karena terlalu banyaknya data dabel dengan akun pembuat yang sama maupun berbeda supaya bisa dihapus dulu data dabel tersebut.

## DAFTAR PUSTAKA

- Darwis, D., Siskawati, N. and Abidin, Z. (2021) 'Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional', *Jurnal Tekno Kompak*, 15(1), p. 131. doi: 10.33365/jtk.v15i1.744.
- Purwiantono, F. E. and Aditya, A. (2020) 'Klasifikasi Sentimen Sara, Hoaks Dan Radikal Pada Postingan Media Sosial Menggunakan Algoritma Naive Bayes Multinomial Text', *Jurnal Tekno Kompak*, 14(2), p. 68. doi: 10.33365/jtk.v14i2.709.
- Rozi, I. F., Hamdana, E. N. and Iqbal Alfahmi, M. B. (2018) 'PENGEMBANGAN APLIKASI ANALISIS SENTIMEN TWITTER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER (Studi Kasus SAMSAT Kota Malang)', *Jurnal Informatika Polinema*, 4(2), p. 149. doi: 10.33795/jip.v4i2.164.
- Samir *et al.* (2021) 'Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes', *Jurnal Media Informatika Budidarma*, 5(1), p. 149. doi: 10.30865/mib.v5i1.2604.