

## ▼ Import Library

```
import tweepy
from textblob import TextBlob
import pandas as pd
import numpy as np
import re,string
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')

from nltk.corpus import stopwords
from wordcloud import WordCloud,STOPWORDS

import csv
```

## Masukan Credential Twitter

```
consumer_key = 'Wn02Vkt660TAWH22cGPHgYVVN'
consumer_secret = 'K000cKr1F8hUYNjd55QEvaAsGkGVerp95J0v1Wq8QWQGt7g9my'
access_token = '2464327753-Q1LiTjp7kXD3xHJlTudQJFvbsEqVRXZj7YkMZaf'
access_token_secret = 'zDc9de9oxYo5JLs5QxnGbVXD9mnPPPg9Kw20NirDoSc5k'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth, wait_on_rate_limit=True)
```

## Crawling untuk mengambil data Indihome

```
csvFile = open('Indihome.csv','a', encoding='utf-8')
csvWriter = csv.writer(csvFile)
for tweet in tweepy.Cursor(api.search, q= "Indihome",count=2000,
                           lang='id').items(2000):
    print(tweet.text)
    csvWriter.writerow([tweet.text])

'''
@IndiHome IndiHome memang selalu memberikan yang terbaik #thnbaruathome
@IndiHome mantap serial bagus nih, mau langsung nonton nanti malem ah 🖐
Banyak tayangan seru bersama indihome #thnbaruathome https://t.co/ew8jyLm2St
@IndiHome Memang indihome yang terbaik deh #thnbaruathome
Indihome tv mantap ada serial yang menarik ceritanya #thnbaruathome https://t.co/S
@IndiHome Eehhh?? Orang tua cuma buat main wa sampe makan speed 15mbps kah?
Indihome TV selalu menemani
#thnbaruathome https://t.co/04oe8nenA6
@IndiHome Indihome tv mantap ada serial yang menarik ceritanya #thnbaruathome
Seru sekali ya indihome #thnbaruathome https://t.co/zb88JwNQ8f
RT @IndiHome: Sebuah kejadian di masa kecilnya membuat Mario (Junior Roberts) memi
@IndiHome Seru sekali ya indihome #thnbaruathome
biar nggak bosan dirumah nonton indihome tv #thnbaruathome https://t.co/1Pe0ZCCbvk
@IndiHome Teman ada film bagus di indihome ini #thnbaruathome
@IndiHome Terima kasih indihome selalu memberikan informasi yang menarik
```

```

#thnb
@IndiHome Yuk sobat indihome kita nonton #thnbaruathome
@IndiHome biar nggak bosan dirumah nonton indihome tv #thnbaruathome
@okkidahlan Halo Kak Okki, ada yang bisa Mumu bantu seputar layanan dan produk Ind
@IndiHome cek dm min
@Fajarra04853240 Hallo Kak Fajar, perihal koneksinya yang kurang maksimal Bondan b
Jangan lupa nonton film ini di indihome tv
#thnbaruathome https://t.co/cnm0609yVR
@IndiHome hiburan dari indihome #thnbaruathome
@IndiHome Progressnya apa ya? Dari kemarin hingga saat ini gak ada progress sama s
Dari sinopsisnya bagus nih drama yang ditayangkan indihome #thnbaruathome https://t.co/NEHC
@IndiHome Jangan lupa nonton film ini di indihome tv
#thnbaruathome
hiburan dari indihome #thnbaruathome https://t.co/LB35Rqhw8O
@IndiHome Wah keren nih indihome #thnbaruathome
@IndiHome Dari sinopsisnya bagus nih indihome #thnbaruathome
Terima kasih IndiHome internet selalu lancar #thnbaruathome https://t.co/Iwyh3cT9t
Langsung cus nonton dong makasih indihome infonya #thnbaruathome https://t.co/NEHC
Indihome memang terbaik
#thnbaruathome https://t.co/e4ezxVr9ju
@IndiHome Sudah benar kata sandinya ka, tiba2 td keputus wifi dihp, trs disambungi
Seru banget pastinya nih IndiHome #thnbaruathome https://t.co/6YdchqxHBq
@IndiHome Wehhh gilaa ini mah flim yang bagus, makasih ya indohome infonya #thnbar
@IndiHome Terima kasih indihome
#thnb
@SanKazes ngsur normal. -Amin(2/2)
@IndiHome wah yang main Junior Roberts sama Zara, seru banget nih pasti
@SanKazes Kak Elza, perihal jaringan layanan tidak pernah melakukan block untuk Ga
Istimewa ini IndiHome
#thnbaruathome https://t.co/GfieVmnKna
RT @IndiHome: Sebuah kejadian di masa kecilnya membuat Mario (Junior Roberts) memi
IndiHome tetap yang terbaik #thnbaruathome https://t.co/5b01uxIDTE
Info film nya bagus terimakasih indihome #thnbaruathome https://t.co/N7ipeq6bnN
Mau nonton apapun juga pakai Indihome pasti lancar #thnbaruathome https://t.co/Qcw
Hiburan terbaik di Indihome #thnbaruathome https://t.co/pOG1dx4GrK
@IndiHome Seru banget pastinya nih IndiHome #thnbaruathome
Seru banget indihome #thnbaruathome https://t.co/rDqdDqhK11
@IndiHome Info film nya bagus terimakasih indihome #thnbaruathome
Seru ya filmnya indihome #thnbaruathome https://t.co/XLtIxLFgE2
mulai kapan nih indihome #thnbaruathome https://t.co/Y7WWDLYYT7
@IndiHome mulai kapan nih indihome #thnbaruathome
@IndiHome Seru ya filmnya indihome #thnbaruathome
@IndiHome Yuk nonton lewat indihome #thnbaruathome

```

```

df = pd.read_csv("Indihome.csv", header=None)
df

```

0

```

0      @joesukasuka Hallo Sobat. Apakah sudah diperti...
1      @prasetyown Hai Kak Prasetyo, maaf sudah bikin...
2      @astagareji Siang, Kak Reji. Silakan pastikan ...
3      @IndiHome DHIR @cvlais TOLONG
4      @IndiHome Silahkan cek DM kak

```

## Lakukan Pra Proses ( Cleaning Data )

```

def praproses(teks):
    teks = re.sub(r'http\S+', '', teks)
    #teks = hapus_tanda(teks)
    teks = re.sub(r'#([\s\S]+)', r'\1', teks) #hapus #tagger
    teks = re.sub(r'@[A-Za-z0-9]+', '', teks) #hapus @
    teks = re.sub(r':([\s\S]+)', r'\1', teks) #hapus #tagger
    teks = re.sub(r'RT[\s\S]+', '', teks) #hapus RT
    teks = re.sub(r'https?:\/\/\S+', '', teks) #hapus hyperlink
    teks = re.sub(r'\w*\d\w*', '', teks).strip() #hapus angka dan angka yang berada dalam st
    teks = hapus_katadouble(teks) #hapus repetisi karakter
    teks = teks.lower() #ubah jadi lower case
    teks = hapus_emoji(teks)
    teks = re.sub(r"[-()\"#/@_%;:+=<>{}=~|.?,]", "", teks)
    teks = re.sub(r"aku", "saya", teks)
    teks = re.sub(r"seller", "penjual", teks)
    teks = re.sub(r"sbg", "sebagai", teks)
    teks = re.sub(r"&amp;", ">", teks)
    return teks

```

```

def hapus_emoji(teks):
    regex_pattern = re.compile(pattern = "["
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        "]+", flags = re.UNICODE)
    return regex_pattern.sub(r'', teks)

```

#Untuk Menghapus kata Double

```

def hapus_katadouble(s):
    pattern = re.compile(r"(\.)\1{1,}", re.DOTALL)
    return pattern.sub(r"\1", s)

```

#Removing the stopwords from text

```

def remove_stopwords(text):
    final_text = []
    for i in text.split():
        if i.strip().lower() not in stop_w:

```

```

    final_text.append(i.strip())
return " ".join(final_text)

```

```

# Load stopwords Bahasa Indonesia
stopword_id = pd.read_csv('stopterm.csv', sep='\t', header=None)
stopword_id.columns = ['word']
stop_w = stopword_id['word'].to_list() #diubah ke list

```

```

#Removing the noisy text
def cleanText(text):
    text = remove_stopwords(text)
    text = praproses(text)
    return text

```

```

#Apply function on review column
df[0] = df[0].apply(cleanText)

```

```
df
```

```

0
0    hallo sobat dipertimbangkan kinanti sarankan d...
1    hai prasetyo maaf bikin gak nyaman kondisi jar...
2    siang reji silakan pastikan terpasang kencang ...
3                                     dhir tolong
4                                     silahkan cek dm
...
3956    nonton film indihome tv thnbaruathome
3957    hiburan indihome thnbaruathome
3958    keren nih indihome thnbaruathome
3959    sinopsisnya bagus nih indihome thnbaruathome
3960    kasih indihome internet lancar thnbaruathome
3961 rows x 1 columns

```

## Tentukan Polarity dan Subjectivity

```

def getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

def getPolarity(text):
    return TextBlob(text).sentiment.polarity

```

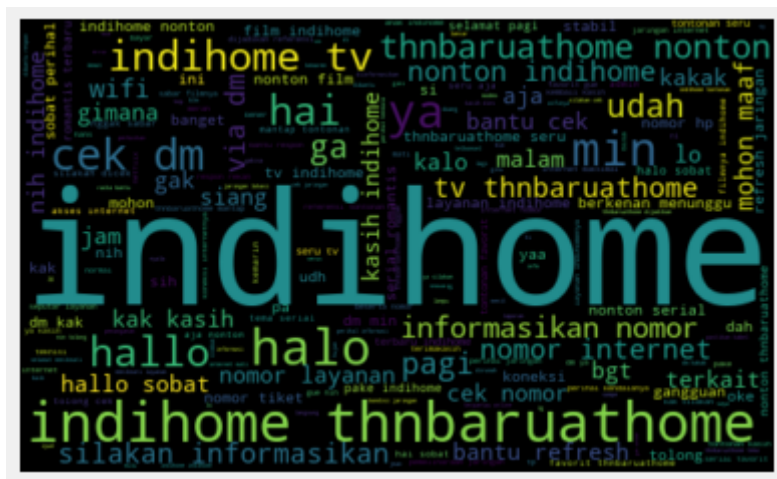
```
df['Polarity'] = df[0].apply(getPolarity)
df['Subjectivity'] = df[0].apply(getSubjectivity)
```

df

		0	Polarity	Subjectivity
0	hallo sobat dipertimbangkan kinanti sarankan d...		0.0	0.0
1	hai prasetyo maaf bikin gak nyaman kondisi jar...		0.0	0.0
2	siang reji silakan pastikan terpasang kencang ...		0.0	0.0
3	dhir tolong		0.0	0.0
4	silahkan cek dm		0.0	0.0
...	...		...	...
3956	nonton film indihome tv thnbaruathome		0.0	0.0
3957	hiburan indihome thnbaruathome		0.0	0.0
3958	keren nih indihome thnbaruathome		0.0	0.0
3959	sinopsisnya bagus nih indihome thnbaruathome		0.0	0.0
3960	kasih indihome internet lancar thnbaruathome		0.0	0.0

3961 rows x 3 columns

```
allWords = ' '.join([twts for twts in df[0]])
wc = WordCloud(width = 500 , height = 300 , random_state=10, max_font_size=110).generate(a
plt.imshow(wc , interpolation = 'bilinear')
plt.axis('off')
plt.show()
```



### Berikan Sentimen pada setiap tweet

```
#untuk menambahkan sentimen positif, negatif / netral dari polarity yg sudah dihitung
def getAnalysisSentiment(score):
    if score < 0:
```

df

	$\theta$	Polarity	Subjectivity	Analysis
0	hallo sobat dipertimbangkan kinanti sarankan d...	0.0	0.0	Neutral
1	hai prasetyo maaf bikin gak nyaman kondisi jar...	0.0	0.0	Neutral
2	siang reji silakan pastikan terpasang kencang ...	0.0	0.0	Neutral
3	dhir tolong	0.0	0.0	Neutral
4	silahkan cek dm	0.0	0.0	Neutral
...	...	...	...	...
3956	nonton film indihome tv thnbaruathome	0.0	0.0	Neutral
3957	hiburan indihome thnbaruathome	0.0	0.0	Neutral
3958	keren nih indihome thnbaruathome	0.0	0.0	Neutral
3959	sinopsisnya bagus nih indihome thnbaruathome	0.0	0.0	Neutral
3960	kasih indihome internet lancar thnbaruathome	0.0	0.0	Neutral

3961 rows x 4 columns

Sesuai data yang ada ternyata analisis sentimen yang didapat yaitu lebih banyak tweet yang netral. untuk mendapat hasil yang lebih baik bisa di lakukan cleaning dan pembobotan kata.Crawling yang saya gunakan yaitu menggunakan bahasa Indonesia. Untuk mendapat hasil yang lebih baik bisa di lakukan cleaning dan pembobotan kata.

Aminurachma Aisyah (17.52.0001)

```
Neutral      3785
Positive     132
Negative      44
Name: Analysis, dtype: int64
```

