

A black and white photograph showing a hand holding a ballot paper over a ballot box. The ballot box is dark and has a slot for the ballot. The background is bright and out of focus.

INDONESIAN 2024 ELECTION SENTIMENT ANALYSIS

Created by
Amanda Daniswara A. (2206051563)
Syazana Alya A. (2206053871)

Pendahuluan

Telah diketahui bahwa pada tahun 2024 akan diselenggarakan pemilihan presiden. Tentunya hal ini menjadi topik hangat bagi segala kalangan di sosial media, terutama di Twitter. Dari beberapa bakal calon presiden kami memilih Prabowo Subianto untuk menjadi topik analisis sentimen. Kami melakukan analisis mengenai topik tersebut dengan menggunakan data yang telah diambil dengan cara crawling Twitter menggunakan module di Python dan Tanpa Twitter API. Data yang kami ambil adalah sebanyak 3205 tweets dengan rentang waktu tanggal 15-17 Juni 2023.

Variabel yang Diperoleh

- | | | | |
|---|----------|----|-----------------|
| 1 | URL | 9 | Conv.ID |
| 2 | Date | 10 | Language |
| 3 | Tweet | 11 | Links |
| 4 | ID | 12 | Media |
| 5 | Replies | 13 | Retweeted Tweet |
| 6 | Retweets | 14 | Bookmarks |
| 7 | Likes | 15 | Username |
| 8 | Quotes | | |

Pre-Processing - 1/3

1 Instalasi dan Import Module

```
[ ] |pip install PySastrawi
import nltk
nltk.download('stopwords')
nltk.download('punkt')
```

```
Collecting PySastrawi
  Downloading PySastrawi-1.2.0-py2.py3-none-any.whl (210 kB)
    210.6/210.6 kB 10.5 MB/s eta 0:00:00
Installing collected packages: PySastrawi
Successfully installed PySastrawi-1.2.0
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
True
```

```
[ ] import pandas as pd
import numpy as np
from datetime import datetime, timedelta
import pytz
import re
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
import ast
import string
from wordcloud import WordCloud
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
import itertools
import matplotlib.pyplot as plt
import seaborn as sns
```

2 Import Data

```
[ ] path = 'https://raw.githubusercontent.com/amandabidah/SentimentAnalysis/main/Twitter%20Data%20-%20Prabowo.csv'
Prabowo = pd.read_csv(path)
Prabowo
```



	URL	Date	Tweet	ID	Replies	Retweets	Likes	Quotes	Conv.	ID
0	https://twitter.com/dikagw2/status/16702197556...	2023-06-17 23:59:55+00:00	Gerindra manggil Prabowo sebagai "The New Soek...	1.670220e+18	0	0	0	0	1.670220e+18	
1	https://twitter.com/novilahamidah16/status/167...	2023-06-17 23:59:51+00:00	Prabowo siap bantu industri pertahanan jadi le...	1.670220e+18	0	0	0	0	1.670220e+18	
2	https://twitter.com/muhhaslim3/status/16702197...	2023-06-17 23:59:48+00:00	Gerindra manggil Prabowo sebagai "The New Soek...	1.670220e+18	0	0	0	0	1.670220e+18	
3	https://twitter.com/miftah_putria/status/16702...	2023-06-17 23:59:39+00:00	Gerindra panggil Prabowo jadi "The New...	1.670220e+18	0	0	0	0	1.670220e+18	

didapat 3205 tweet dengan rentang waktu 15 - 17 Juni 2023

Pre-Processing - 2/3

3

Folding

```
Prabowo = Prabowo.drop(['URL', 'Replies', 'Retweets',  
                        'Likes', 'Quotes', 'Links', 'Media',  
                        'ID', 'Retweeted Tweet', 'Bookmarks',  
                        'Conv. ID', 'Language', 'Username',  
                        'Date'], axis=1)  
  
Prabowo.head()
```



Tweet

```
0 Gerindra manggil Prabowo sebagai "The New Soek...  
1   Prabowo siap bantu industri pertahanan jadi le...  
2 Gerindra manggil Prabowo sebagai "The New Soek...  
3 Gerindra panggil Prabowo jadi "The New Soekarn...  
4   Prabowo: Industri pertahanan harus terus berke...
```

```
def MembersihkanText(Tweet):  
    Tweet = re.sub(r'@[A-Za-z0-9]+', '', Tweet)  
    Tweet = re.sub(r'#[A-Za-z0-9]+', '', Tweet)  
    Tweet = re.sub(r'RT[\s]', '', Tweet)  
    Tweet = re.sub(r'http\S+', '', Tweet)  
    Tweet = re.sub(r'[0-9]+', '', Tweet)  
    Tweet = Tweet.replace('\n', ' ')  
    Tweet = Tweet.translate(str.maketrans('', '', string.punctuation))  
    Tweet = Tweet.strip(' ')  
    return Tweet  
  
def LowercaseText(text):  
    text = text.lower()  
    return text
```

Pre-Processing - 3/3

4 Tokenizing

```
def TokenisasiText(text):  
    text = word_tokenize(text)  
    return text
```

5 Filtering

```
def FilterText(text):  
    listStopwords = set(stopwords.words('indonesian'))  
    filtered = []  
    for txt in text:  
        if txt not in listStopwords:  
            filtered.append(txt)  
    text = filtered  
    return text
```

5 Filtering

```
def StemmingText(text):  
    factory = StemmerFactory()  
    stemmer = factory.create_stemmer()  
    text = [stemmer.stem(word) for word in text]  
    return text  
  
def Kalimat(list_words):  
    sentence = ' '.join(word for word in list_words)  
    return sentence
```

```
Prabowo['Tweet_Cleaned'] = Prabowo['Tweet'].apply(MembersihkanText)  
Prabowo['Tweet_Cleaned'] = Prabowo['Tweet_Cleaned'].apply(LowercaseText)  
Prabowo['Tweet_Cleaned'] = Prabowo['Tweet_Cleaned'].apply(TokenisasiText)  
Prabowo['Tweet_Cleaned'] = Prabowo['Tweet_Cleaned'].apply(FilterText)  
Prabowo['Tweet_Processed'] = Prabowo['Tweet_Cleaned'].apply(StemmingText)  
Prabowo['Tweet_Cleaned'] = Prabowo['Tweet_Cleaned'].apply(Kalimat)
```

```
Prabowo.drop_duplicates(subset = 'Tweet', inplace = True)  
NAMA_FILE_CSV = 'Prabowo_Pre-Processed.csv'  
Prabowo.to_csv(NAMA_FILE_CSV, index=False)
```

Hasil Pre-Processing



	Tweet	Tweet_Cleaned	Tweet_Processed
0	Gerindra manggil Prabowo sebagai "The New Soek...	gerindra manggil prabowo the new soekarno habi...	[gerindra, manggil, prabowo, the, new, soekarn...
1	Prabowo siap bantu industri pertahanan jadi le...	prabowo bantu industri pertahanan mandiri koal...	[prabowo, bantu, industri, tahan, mandiri, koa...
3	Gerindra panggil Prabowo jadi "The New Soekarn...	gerindra panggil prabowo the new soekarno tawa...	[gerindra, panggil, prabowo, the, new, soekarn...
4	Prabowo: Industri pertahanan harus terus berke...	prabowo industri pertahanan berkembang koalisi...	[prabowo, industri, tahan, kembang, koalisiskeb...
7	Gerindra sebut Prabowo kayak "The New Soekarno...	gerindra prabowo kayak the new soekarno garaga...	[gerindra, prabowo, kayak, the, new, soekarno,...
...
3202	"China dan Indonesia harus memainkan peran yan...	china indonesia memainkan peran signifikan men...	[china, indonesia, main, peran, signifikan, ja...
3203	Wei juga menekankan bahwa situasi saat ini di ...	wei menekankan situasi kawasan asiapasifik men...	[wei, tekan, situasi, kawasan, asiapasifik, ha...
3204	ekonomi, budaya, dan maritim, yang secara jela...	ekonomi budaya maritim arah bersamasama memban...	[ekonomi, budaya, maritim, arah, bersamasama, ...
3205	Dalam keterangan resmi Kementerian Pertahanan ...	keterangan resmi kementerian pertahanan china ...	[terang, resmi, menteri, tahan, china, wei, me...
3206	Menhan Prabowo Kunjungi China, Mau Pacu Kerja ...	menhan prabowo kunjungi china pacu kerja milit...	[menhan, prabowo, kunjung, china, pacu, kerja,...

3002 rows × 3 columns

Analisis Sentimen

```
lexicon_positive = dict()
import csv
with open('/content/lexicon_positive.csv', 'r') as csvfile:
    reader = csv.reader(csvfile, delimiter=',')
    for row in reader:
        lexicon_positive[row[0]] = int(row[1])

lexicon_negative = dict()
import csv
with open('/content/lexicon_negative.csv', 'r') as csvfile:
    reader = csv.reader(csvfile, delimiter=',')
    for row in reader:
        lexicon_negative[row[0]] = int(row[1])

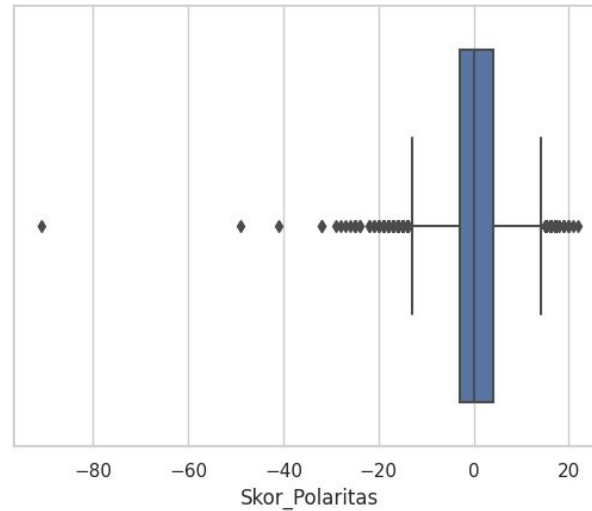
def AnalisisSentimen(text):
    score = 0
    for word in text:
        if (word in lexicon_positive):
            score = score + lexicon_positive[word]
    for word in text:
        if (word in lexicon_negative):
            score = score + lexicon_negative[word]
    Polaritas=''
    if (score > 0):
        Polaritas = 'positive'
    elif (score < 0):
        Polaritas = 'negative'
    else:
        Polaritas = 'neutral'
    return score, Polaritas
```

Menggunakan word dictionary yang bersumber dari <https://github.com/fajri91/InSet>

- skor polaritas <0 : negatif
- skor polaritas = 0 : netral
- skor polaritas >0 : positif

Persebaran sentimen memiliki nilai median 0 yang menandakan persebaran sentimen yang tidak merata

Box-plot persebaran nilai sentimen



dengan boxplot, didapat median bernilai 0 dan persebaran nilai tidak merata.

Kesimpulan

Kita dapat menyimpulkan bahwa:

- 1 Sentimen pengguna twitter terhadap prabowo mayoritas bersifat positif.
- 2 Kata yang paling banyak digunakan dalam tweet adalah "prabowo" kemudian disusul oleh "gerindra" dan "soekarno".
- 3 Persebaran polaritas tweet tidak terlalu merata karena terdapat tweet yang nilainya sangat negatif.



Terima Kasih

Created by
Amanda Daniswara A. (2206051563)
Syazana Alya A. (2206053871)