

API FOR TEXT CLEANSING AND DATA ANALYSIS REPORTS

Indonesian Abusive and Hate Speech Twitter Text Analysis

By
Aldimeola Alfarisy

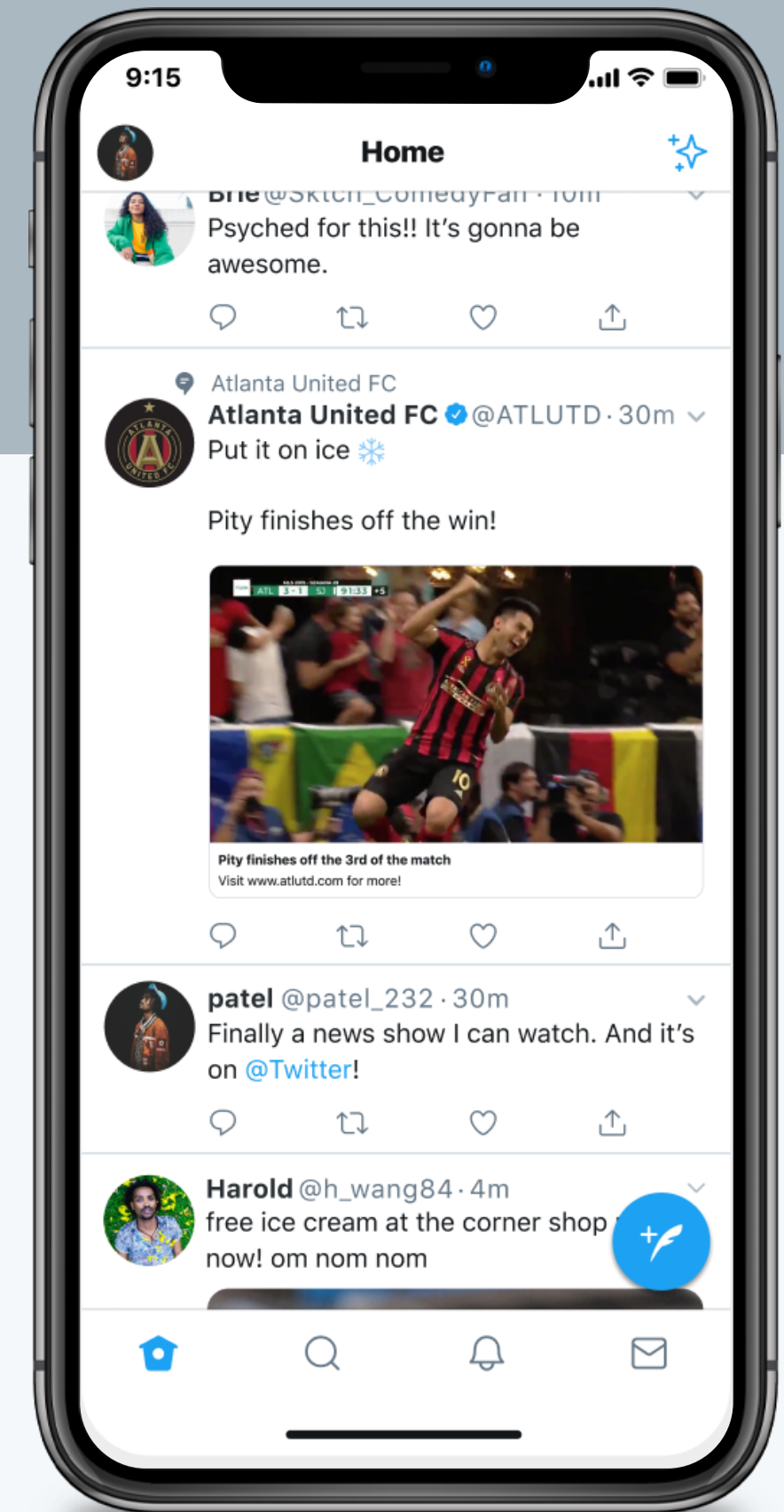


Table of Contents



Pendahuluan



Metode Penelitian



Hasil dan Kesimpulan

Pendahuluan

Latar Belakang

Ujaran kebencian adalah suatu bentuk ekspresi yang dilakukan untuk menghilangkan kebencian dan melakukan tindakan kekerasan serta menentang seseorang atau sekelompok orang dengan berbagai alasan. Kasus ujaran kebencian sangat sering ditemui di media sosial, salah satunya di Twitter.

Dalam laporan berjudul 'Digital Civility Index (DCI)' yang diterbitkan oleh Microsoft pada tahun 2020, Indonesia menempati peringkat terendah se-Asia Tenggara untuk tingkat kesopanan atau peringkat ke-29 dari 32 negara yang disurvei. Sedangkan menurut data POLRI, sejak April 2020 hingga Juli 2021 tercatat ada 937 kasus yang dilaporkan dengan jumlah terbanyak yaitu kasus provokatif, konten kebencian dan ujaran kebencian.

Berdasarkan hal tersebut, saya akan menganalisis apakah tweet tersebut mengandung ujaran kebencian atau kata-kata kasar

Rumusan Masalah

- Berapa banyak jumlah atau perbandingan tweet yang mengandung kata-kata kasar dan ujaran kebencian dari keseluruhan tweet yang ada?
- Kepada siapa tweet ujaran kebencian tersebut dituju?
- Topik apa yang paling sering dibicarakan dan bagaimana tingkatan ujaran kebencian pada tweet?
- Bagaimana caranya agar lebih mudah memahami maksud dari sebuah tweet?

Tujuan

- Mengetahui jumlah, tujuan, serta terkait apa pesan-pesan negatif yang terkandung pada tweet yang ditulis oleh orang-orang Indonesia.
- Memproses text pada tweet agar kata-kata yang digunakan lebih baku.

Metode Penelitian

Deskripsi Data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13169 entries, 0 to 13168
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Tweet                 13169 non-null  object
1   HS                    13169 non-null  int64
2   Abusive               13169 non-null  int64
3   HS_Individual         13169 non-null  int64
4   HS_Group              13169 non-null  int64
5   HS_Religion           13169 non-null  int64
6   HS_Race               13169 non-null  int64
7   HS_Physical           13169 non-null  int64
8   HS_Gender             13169 non-null  int64
9   HS_Other              13169 non-null  int64
10  HS_Weak               13169 non-null  int64
11  HS_Moderate           13169 non-null  int64
12  HS_Strong             13169 non-null  int64
dtypes: int64(12), object(1)
memory usage: 1.3+ MB
```

```
In [5]: # Duplicated data check
data.duplicated().sum()
```

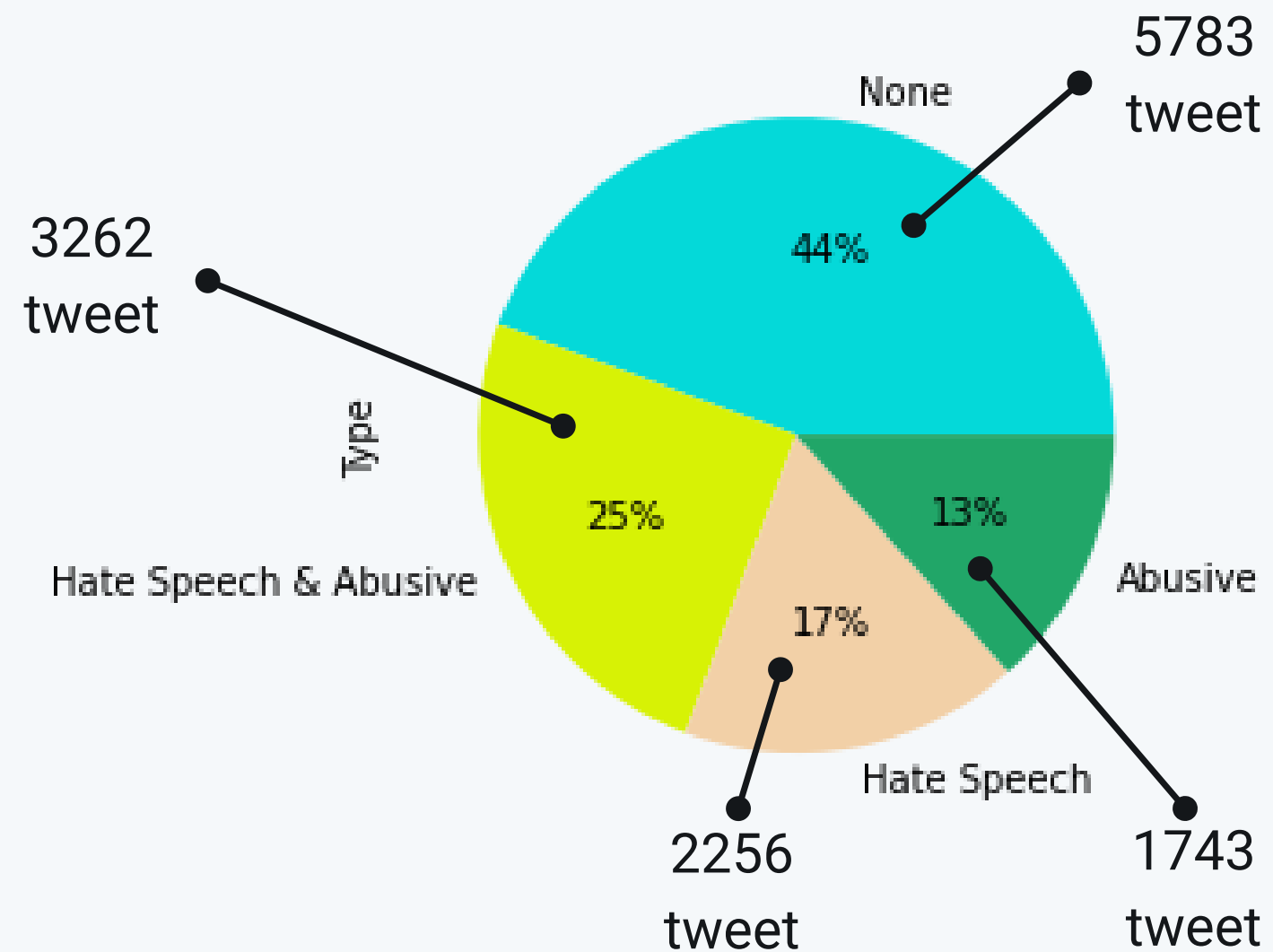
```
Out[5]: 125
```

- Dataset diperoleh dari kaggle
- Dataset terdiri 13,169 row dan 13 column yang berisikan tweet dalam bahasa Indonesia serta klasifikasi nya
- Tidak terdapat missing values dalam dataset
- Terdapat 125 duplikat data dan duplikat data sudah di hilangkan (dari 13,169 data menjadi 13,044 data)

Metode Penelitian

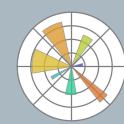
Analisis Data

Perbandingan Tipe Tweet

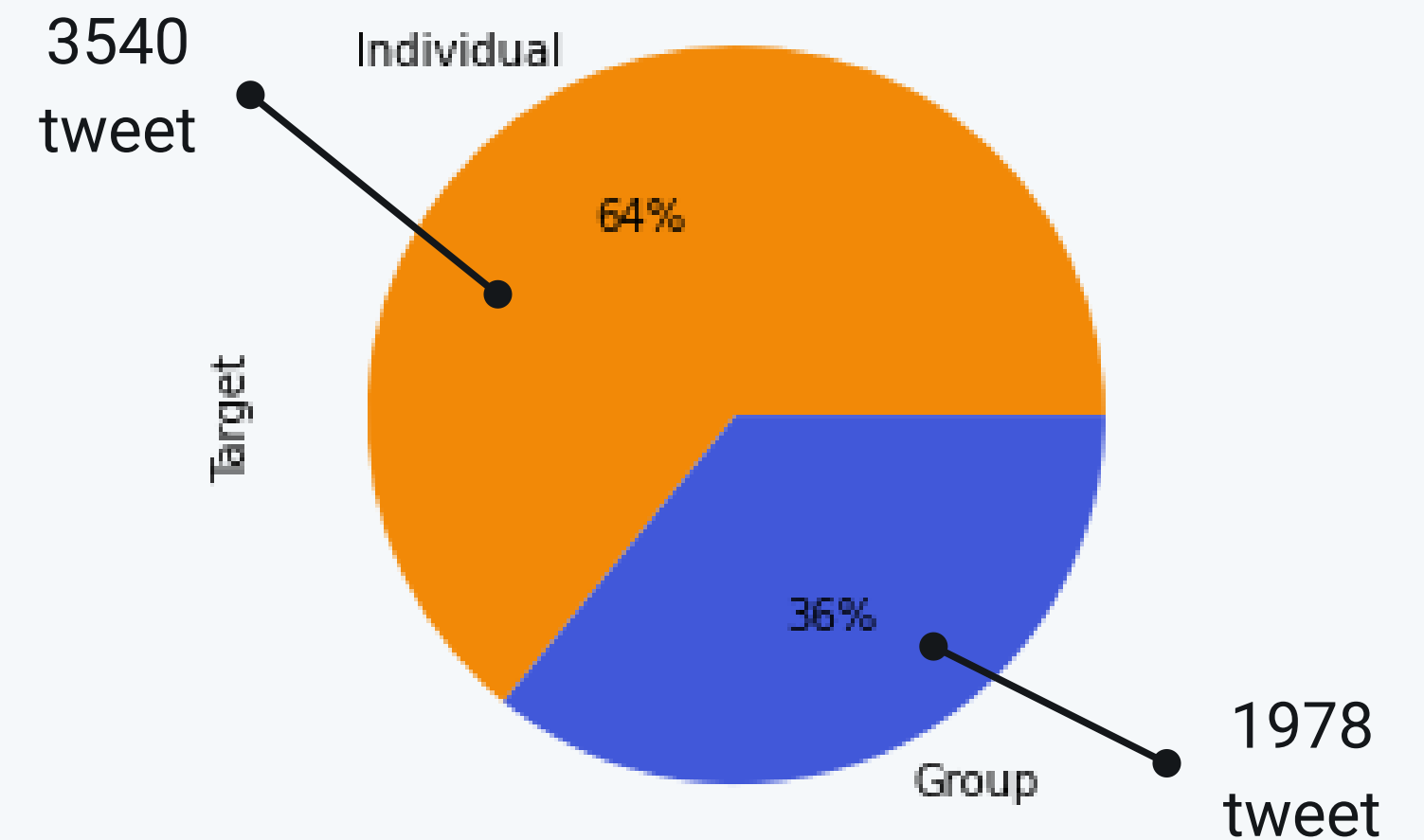


Terdapat 7261 (56%) tweet negatif

Tools and Libraries :



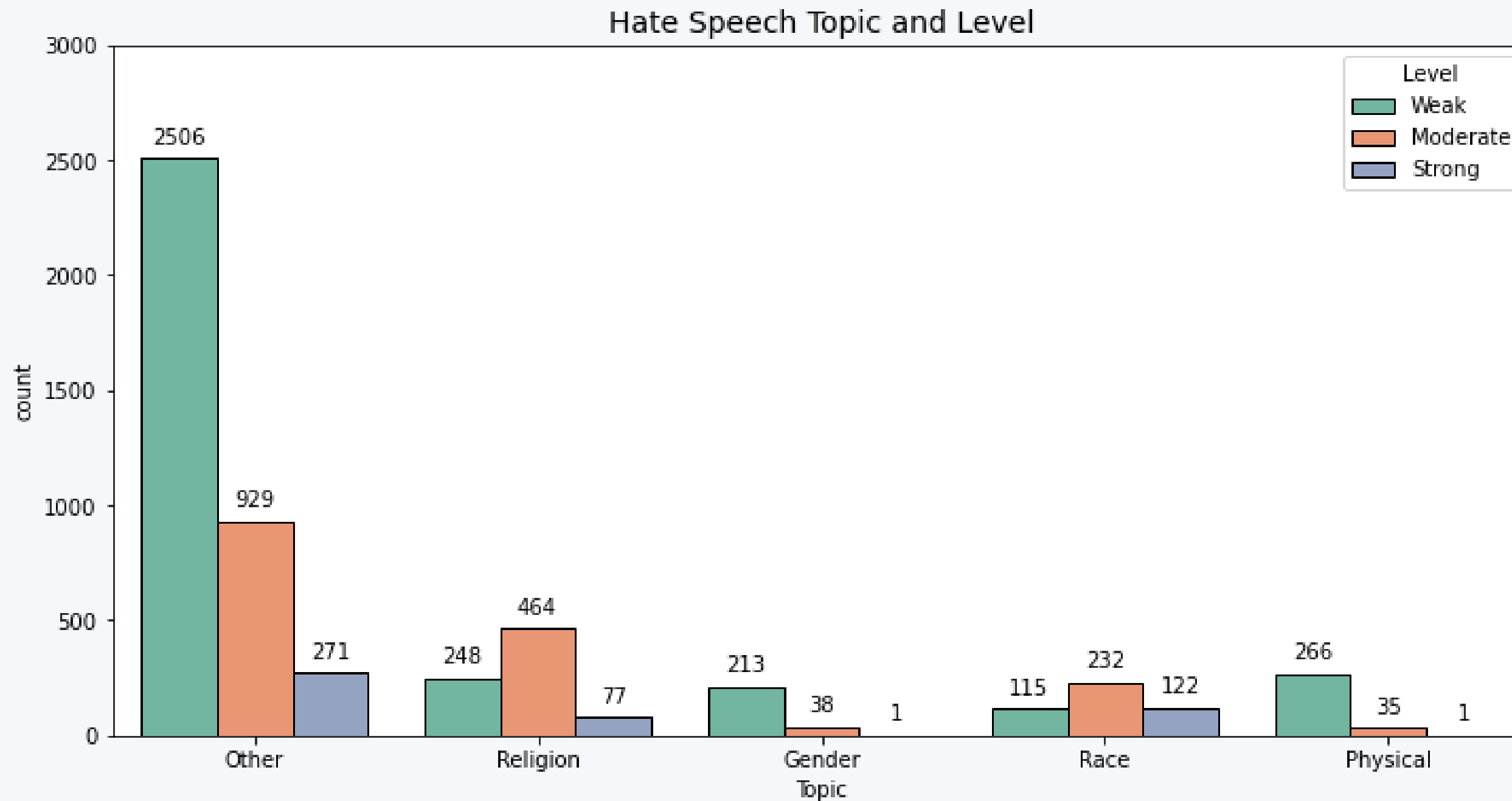
Perbandingan Target Tweet



Total 5518 tweet mengandung hate speech

Metode Penelitian

Analisis Data



Topik ujaran kebencian yang sering diperbincangkan adalah topik yang general (Other), diikuti dengan dengan topik agama. Begitu pula dari tingkat ujaran kebencian yang dibahas mayoritas pada topik general (Other) dan diikuti topik agama.

Tools and Libraries :



Metode Penelitian

Text Processing

Text Cleansing

```
def lowercase(s):
    return s.lower()

def remove_punctutation(s):
    s = re.sub('[^0-9a-zA-Z]+', ' ', s)
    s = re.sub(r':', ' ', s)
    s = re.sub('\n', ' ', s)
    s = re.sub('rt', ' ', s)
    s = re.sub('((www\.[^\s]+)|(https?://[^\s]+)|(http?://[^\s]+))', ' ', s)
    s = re.sub(' +', ' ', s)
    s = re.sub(r'pic.twitter.com.[\w]+', ' ', s)
    s = re.sub('user', ' ', s)
    s = re.sub('gue', 'saya', s)
    s = re.sub(r',\s', ' ', s)
    return s

db = sqlite3.connect('C:/Users/ASUS/Documents/Gold Challenge/database.db', check_same_thread = False)
q_kamusalay = 'SELECT * FROM kamusalay'
t_kamusalay = pd.read_sql_query(q_kamusalay, db)
q_abusive = 'SELECT * FROM abusive'
t_abusive = pd.read_sql_query(q_abusive, db)

alay_dict = dict(zip(t_kamusalay['alay'], t_kamusalay['normal']))
def alay_to_normal(s):
    for word in alay_dict:
        return ' '.join([alay_dict[word] if word in alay_dict else word for word in s.split(' ')])

l_abusive = t_abusive['ABUSIVE'].str.lower().tolist()
def normalize_abusive(s):
    list_word = s.split()
    return ' '.join([s for s in list_word if s not in l_abusive])
```

Mengubah karakter menjadi lowercase

Menghilangkan karakter yang tidak diperlukan

Mengganti kata-kata yang tidak baku menjadi kata yang baku

Menghilangkan kata-kata abusive

Tools and Libraries :



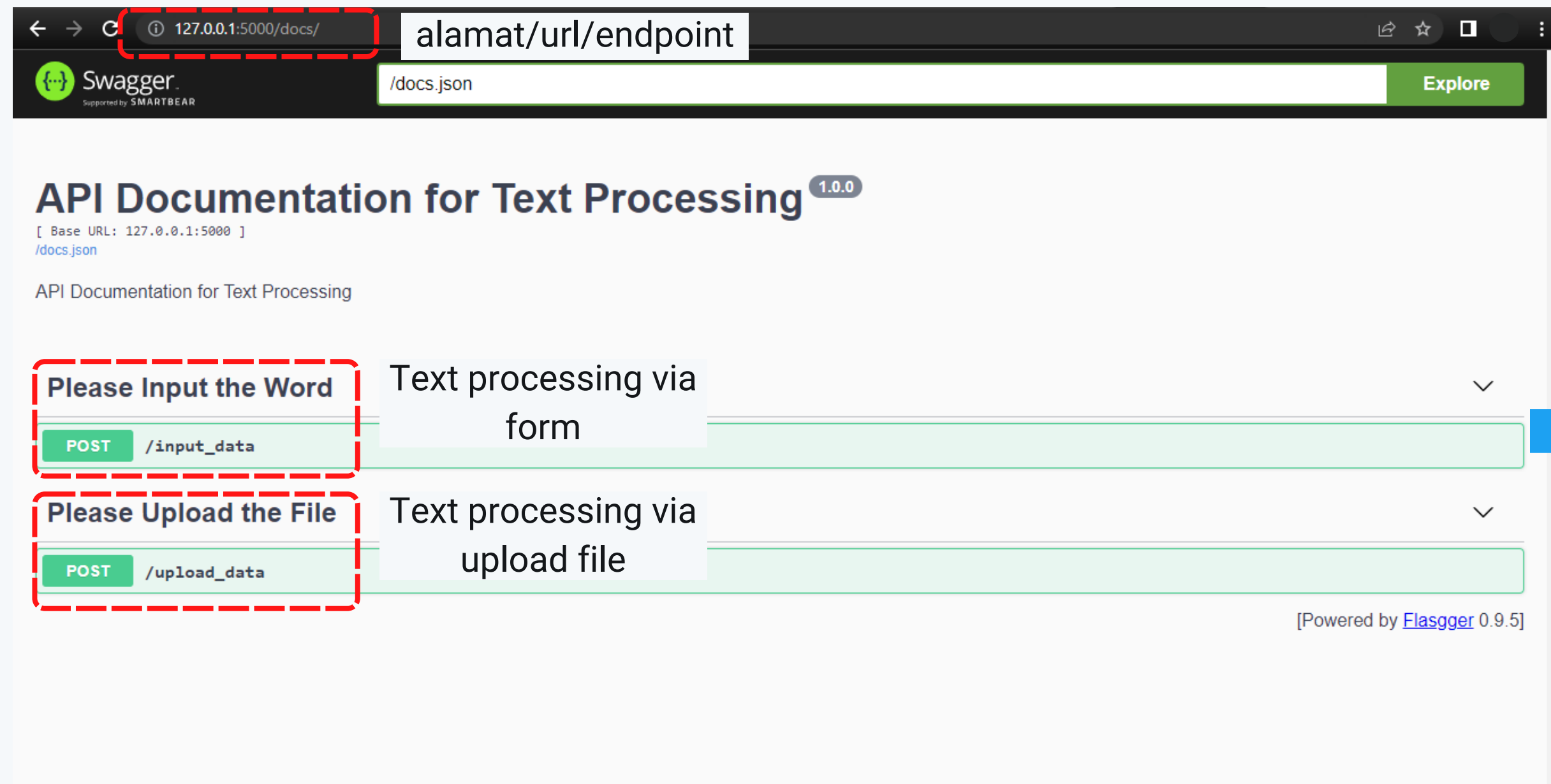
RegEx



Metode Penelitian

Text Processing

API (Application Program Interface) Text Processing



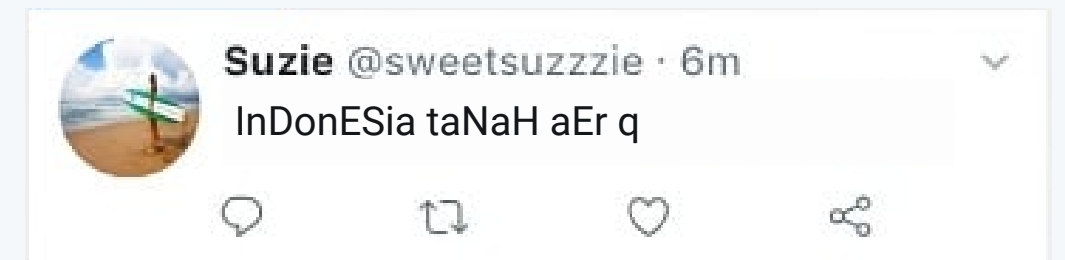
The screenshot shows the Swagger API documentation for a text processing service. The browser address bar displays the URL `127.0.0.1:5000/docs/alamat/url/endpoint`. The Swagger interface includes a search bar with `/docs.json` and an `Explore` button. The main heading is **API Documentation for Text Processing 1.0.0**, with a base URL of `127.0.0.1:5000`. Two API endpoints are listed:

- Please Input the Word**: A POST endpoint at `/input_data` for "Text processing via form".
- Please Upload the File**: A POST endpoint at `/upload_data` for "Text processing via upload file".

Both endpoint descriptions are highlighted with red dashed boxes. A blue arrow points from the first endpoint to the right, indicating the flow of the example.

Contoh Input & Output Teks

Tweet lama:



Tweet baru:



Tools and Libraries :



Hasil dan Kesimpulan

Hasil

- Dari 13044 kumpulan tweet, terdapat 7261 (56%) tweet negatif dengan komposisi 1743 (13%) tweet mengandung kata-kata kasar, 2256 (17%) tweet mengandung ujaran kebencian, dan 3262 (25%) tweet mengandung keduanya
- Dari 5518 tweet negatif yang mengandung hate speech dan abusive Mayoritas tweet dengan ujaran kebencian ditujukan kepada individu tertentu sejumlah 3540 (64%) tweet
- Topik ujaran kebencian yang sering diperbincangkan adalah topik yang general (Other). Begitu pula dengan tingkat ujaran kebencian yang dibahas mayoritas pada topik general (Other)
- Salah satu cara agar data teks seperti tweet lebih mudah dipahami maksudnya adalah dengan di processing (cleansing) dengan bantuan library ReGex. Lalu untuk sekumpulan tweet dalam jumlah banyak yang digabung dalam suatu data, dapat di cleansing secara relatif singkat dengan membuat API (Application Program Interface) menggunakan library Flask serta Swagger

Kesimpulan dan Saran

Berdasarkan hasil analisa yang telah dilakukan, bisa kita simpulkan bahwa pengguna social media (khususnya twitter) para warga Indonesia mayoritas membuat tweet yang mengandung kata-kata kasar serta ujaran kebencian. Kita dapat mengatur penggunaan media sosial menjadi lebih positif, salah satunya dengan tidak mengikuti akun-akun yang dapat memicu kebencian dengan kumpulan tweet nya yg provokatif. Jika dirasa perlu, melaporkan tweet atau akun tersebut kepada pihak aplikasi (twitter) dapat dilakukan sehingga akun tersebut dapat dihilangkan.

Dokumentasi



github.com/aldimeolaalfarisy/Indonesian-Abusive-and-Hate-Speech-Twitter-Text-Analysis

TERIMA KASIH

Contact



0878-7375-6695



aldimeolaalfarisy@yahoo.com



aldimeolaa