



Hate Tweet Analysis

Imaduddin Abdul Karim

GOLD CHALLENGE

PENDAHULUAN

Hate Tweet Analysis

Latar Belakang

Peningkatan penggunaan media sosial turut diiringi dengan meningkatnya fenomena hate speech, yang dapat menyebabkan dampak negatif secara sosial dan psikologis. Penelitian sebelumnya menunjukkan adanya rentan kelompok atau individu menjadi target utama hate speech. Oleh karena itu, perlu dilakukan analisis untuk memahami target hate speech, tingkat intensitasnya, dan pola penggunaan kata-kata abusive yang terlibat.

Penelitian ini bertujuan untuk mengidentifikasi pola hate speech di media sosial, fokus pada kelompok atau individu yang rentan, dan menganalisis korelasi antara jumlah kata abusive dengan intensitas hate speech. Dengan pemahaman mendalam terhadap fenomena ini, diharapkan dapat dikembangkan strategi mitigasi yang lebih efektif untuk menciptakan lingkungan online yang lebih aman dan inklusif.

Tujuan

- Mengidentifikasi apakah kelompok ataukah individu yang paling sering dijadikan target hate speech beserta tingkat hate speech yang diterima.
- Menentukan topik yang paling sering menjadi sasaran hate speech dan menganalisis tingkat keparahan hate speech terhadap topik tersebut.
- Meneliti korelasi antara jumlah kata abusive dengan banyaknya kata yang di-tweet seseorang.
- Mengidentifikasi kata-kata abusive yang paling sering digunakan dalam tweet dan di setiap topik.

1

2

3

4

Rumusan Masalah

- Siapa yang paling sering dijadikan target hate speech dan bagaimana tingkat hate speech yang diterima?
- Topik apa yang paling sering digunakan untuk menjadi hate speech dan bagaimana tingkat hate speech terhadap topik tsb?
- Apakah terdapat korelasi antara jumlah kata abusive dengan banyaknya kata?
- Kata Abusive apa saja yang paling banyak digunakan dalam data Tweet dan di setiap topik?

Batasan Masalah

- Penelitian ini terfokus pada hate speech yang terjadi di twitter dari dataset [Indonesian Abusive and Hate Speech Twitter Text](#)
- Lingkup target hate speech mencakup kelompok atau individu berdasarkan topik ras, agama, gender, dan fisik.
- Tingkat hate speech didapat dari dataset yang memiliki tingkatan Weak, Moderate, dan Strong
- Kata-kata abusive yang dianalisis terbatas pada kategori tertentu yang mencakup frasa atau kata-kata yang dianggap merendahkan dan merugikan.

Metode Penelitian

Deskripsi Data

#	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
13	Tweet_Cleaned	13169 non-null	object

1

Data Sekunder

Data yang digunakan pada penelitian ini adalah data sekunder yang didapat dari [Indonesian Abusive and Hate Speech Twitter Text](#)

2

13.044 Data

Data berjumlah 13.169 baris dan 13 kolom, sebelum menghilangkan data duplikat. Total data yang tidak duplikat adalah 13.044.

3

Tweet_Cleaned

Kolom Tweet_Cleaned merupakan hasil dari cleansing data kolom Tweet

4

Data lain

Terdapat juga beberapa data seperti abusive, dan kamusalay yang akan membantu penelitian ini

METODE PENELITIAN

PERSIAPAN DATABASE SQLITE3

stopword_table

Tabel ini berisi kata-kata yang termasuk dalam kategori stopwords, bersumber dari library Sastrawi. Data ini akan digunakan pada proses cleansing data.

abusive_table

Tabel ini berisi data yang diinputkan ke challenge_database.db yang bersumber dari abusive.csv. Data ini berupa kata-kata abusive yang akan digunakan untuk pada tahap analisis.



kamus_alay_table

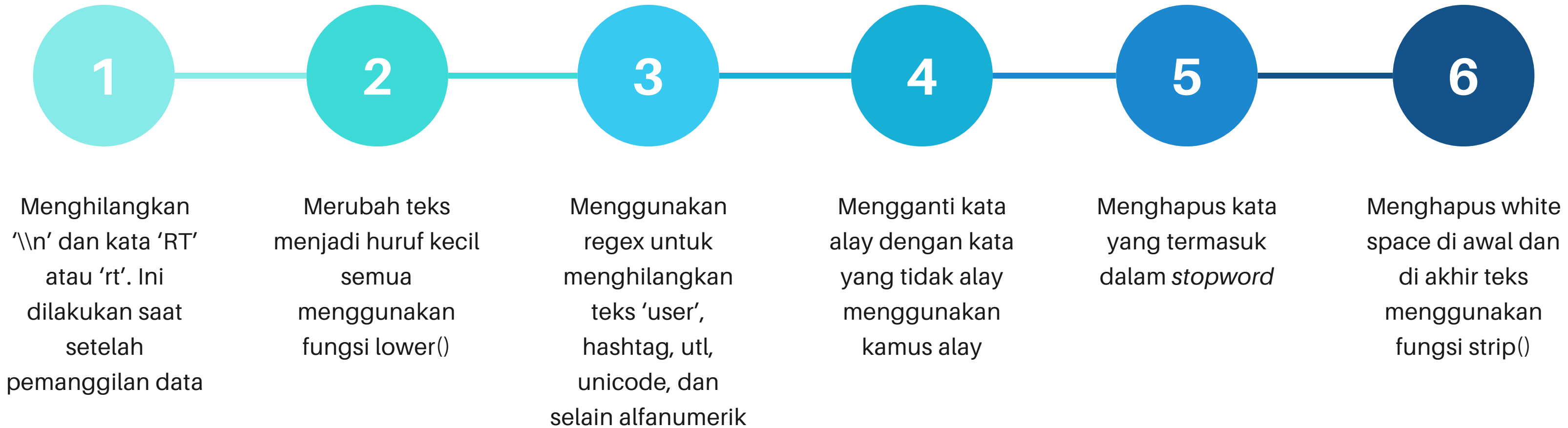
Tabel ini berisi kamus kosa kata alay yang bersumber dari new_kamusalay.csv. Data ini memiliki 2 kolom yaitu "alay" dan "tidak_alay" yang akan digunakan pada proses cleansing data.

tweet_table

Seluruh data di dalam file data.csv akan di input ke dalam database dengan nama tabel tweet_table, namun proses input ini akan dilakukan bersamaan dengan hasil API cleansing data.

METODE PENELITIAN

CLEANSING DATA



	Tweet	Tweet_Cleaned
0	- disaat semua cowok berusaha melacak perhatia...	semua berusaha melacak gue kamu lantas remehka...
1	RT USER: USER siapa yang telat ngasih tau elu?...	telat memberi tau kamu sarap gue bergaul cigax...
2	41. Kadang aku berfikir, kenapa aku tetap perc...	41 aku berpikir aku tetap tuhan aku selalu jat...
3	USER USER AKU ITU AKU\n\nKU TAU MATAMU SIPIT T...	aku aku ku tau matamu sipit mana aku
4	USER USER Kaum cebong kapir udah keliatan dong...	kaum cebong kafir awal tambah dungu haha

METODE PENELITIAN

API CONFIGURATION

POST /clean_teks

Endpoint ini memerlukan input teks dan beberapa parameter tambahan untuk diproses.

POST /clean_teks

Parameters

Name	Description
text * required string (formData)	RT USER AHhh Dasar BABU LU?!
hs * required integer (query)	hs 1
abusive * required integer (query)	abusive 1
target * required string (query)	target HS_Individual
topic * required string (query)	topic HS_Other
level * required string (query)	level HS_Weak

Response body

```
{
  "Abusive": 1,
  "HS": 1,
  "HS_Gender": 0,
  "HS_Group": 0,
  "HS_Individual": 1,
  "HS_Moderate": 0,
  "HS_Other": 1,
  "HS_Physical": 0,
  "HS_Race": 0,
  "HS_Religion": 0,
  "HS_Strong": 0,
  "HS_Weak": 1,
  "Tweet": "RT USER AHhh Dasar BABU LU?!",
  "Tweet_Cleaned": "ah dasar babu kamu"
}
```

POST /text-processing-file

Endpoint ini memerlukan input file CSV dengan kolom tertentu untuk diproses. Kolom-kolom tersebut telah dijelaskan pada slide metode penelitian - deskripsi data.

POST /text-processing-file

Parameters

Name	Description
file * required file (formData)	File to upload Choose file data.csv

Response body

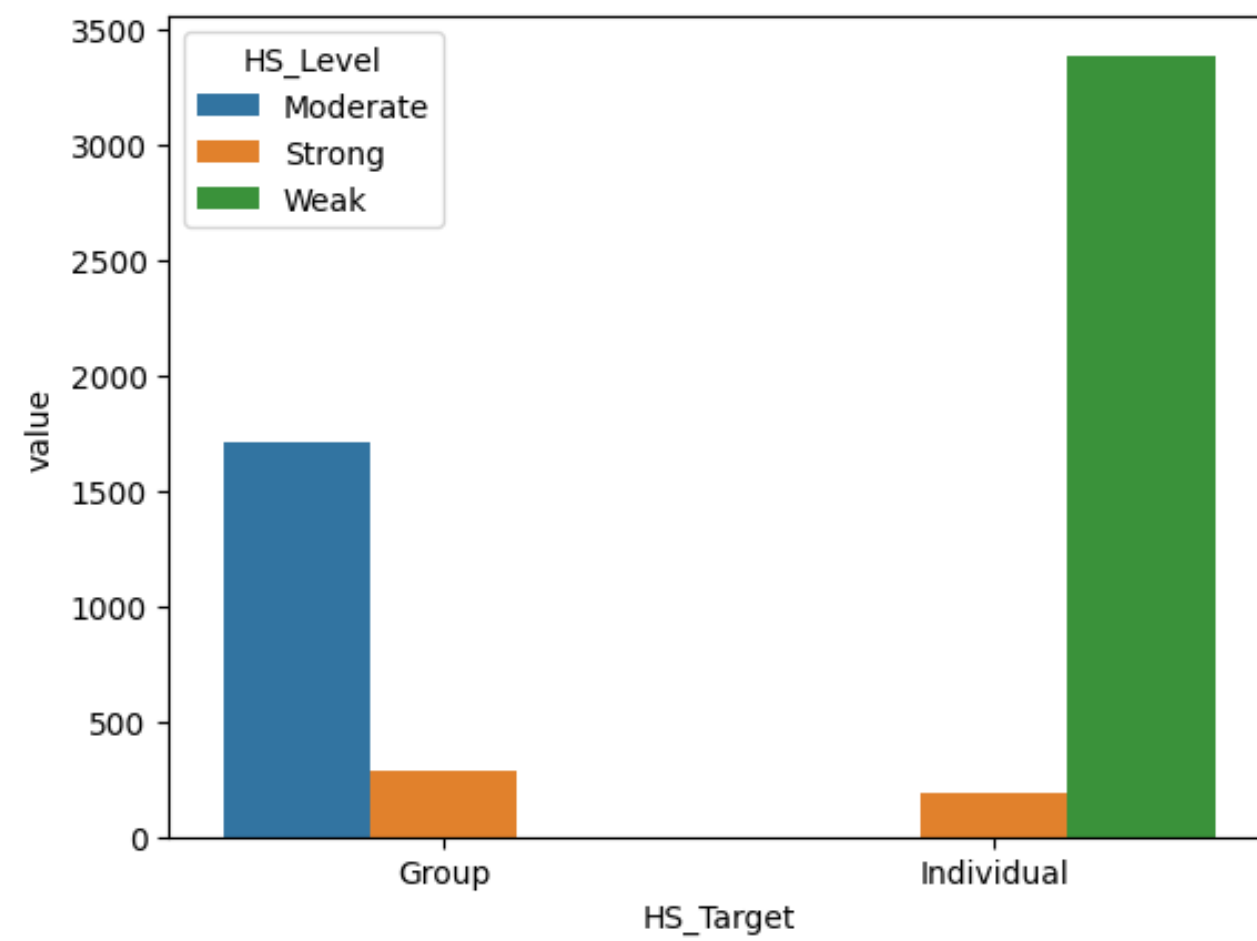
```
{
  "Tweet_Cleaned": [
    "semua cowok berusaha melacak perhatian",
    "siapa telat memberi tau kamu edan sarap",
    "41 kadang aku berpikir aku tetap percaya",
    "memilih jadi kristen aku anak ter",
    "aku aku ku tau matamu sipit dilihat man",
    "kaum cebong kafir kelihatan dongoknya a",
    "bani taplak kawan kawan",
    "deklarasi pilihan kepala daerah 2018 am",
    "gue baru selesai re watch aldnoah zero",
    "nah admin belanja satu port terbaik nak",
    "senawang",
    "enak kalau ngewe",
    "gue punya jari tengah buat kamu gue ukur",
    "banci kaleng malu jawab pertanyaan 2 ha",
    "kalau belajar ekonomi mestinya jago mem",
    "aktor huru hara 98 prabowo si lengserka",
    "bu guru enakan jadi jablay guru sekolah",
    "lawan bicara gue intelek kayak kamu ota",
    "belakangan kok pikiran ampas banget",
    "ari sama beki rapi monyet",
    "..."
  ]
}
```

HASIL DAN KESIMPULAN

VISUALISASI DATA

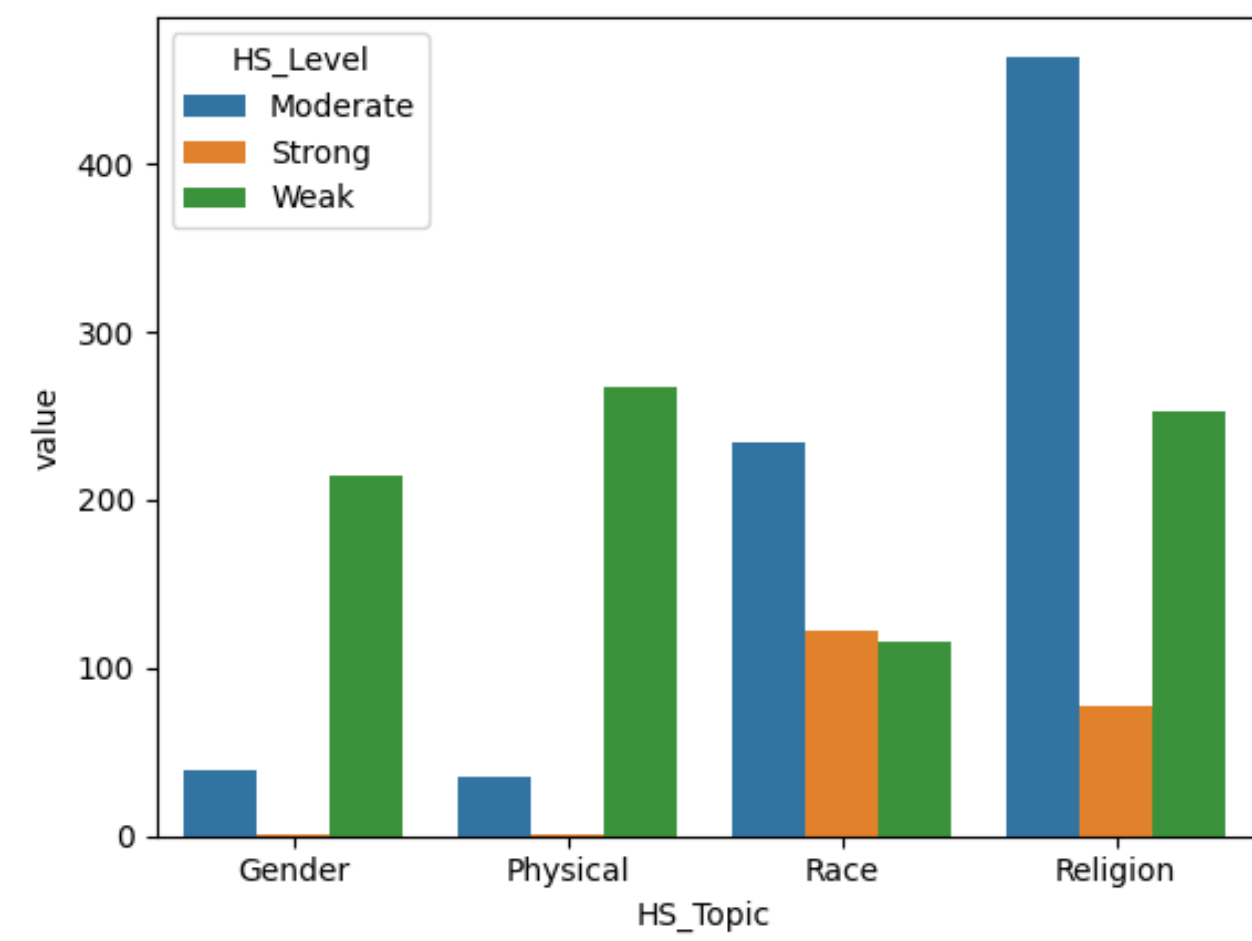
Hate Speech Level terhadap Target

Target **Group** menjadi yang terbanyak dalam menerima hate speech di level **moderate** dan **strong** dibandingkan individual



Hate Speech Level terhadap Topic

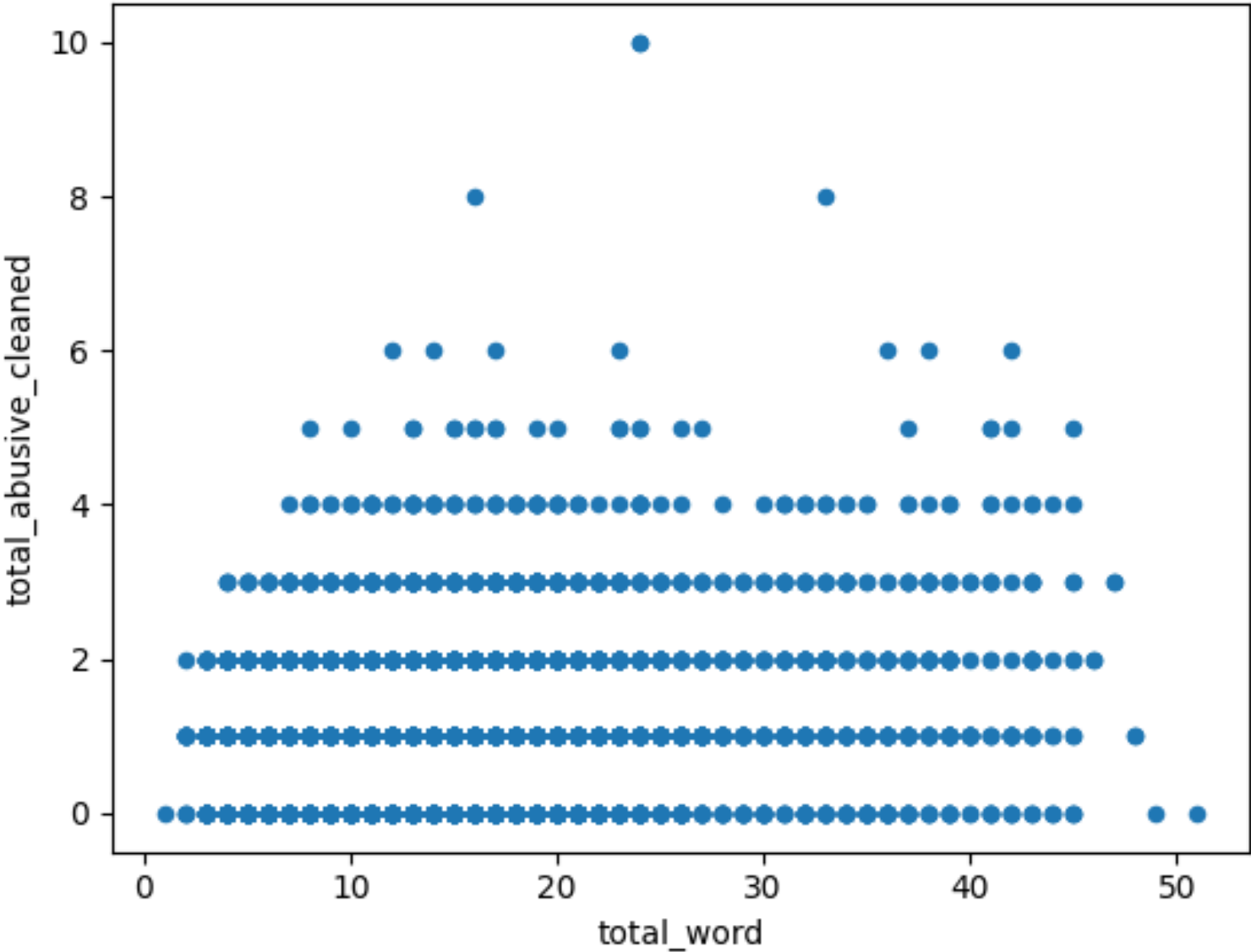
Topik **Religion** menjadi topik yang paling banyak dibahas sekaligus yang terbanyak di level **Moderate**. Namun topik **Race** menjadi yang terbanyak di level **Strong**.



	total_char	total_word	total_char_cleaned	total_word_cleaned	total_abusive_cleaned
total_char	1.000000	0.948519	0.842500	0.845408	0.159032
total_word	0.948519	1.000000	0.853101	0.875217	0.144450
total_char_cleaned	0.842500	0.853101	1.000000	0.978819	0.191674
total_word_cleaned	0.845408	0.875217	0.978819	1.000000	0.208692
total_abusive_cleaned	0.159032	0.144450	0.191674	0.208692	1.000000

HASIL DAN KESIMPULAN

VISUALISASI DATA



Tidak terdapat kolerasi antara jumlah kata abusive terhadap jumlah kata yang terdapat pada *hate speech tweet*.

WORDCLOUD



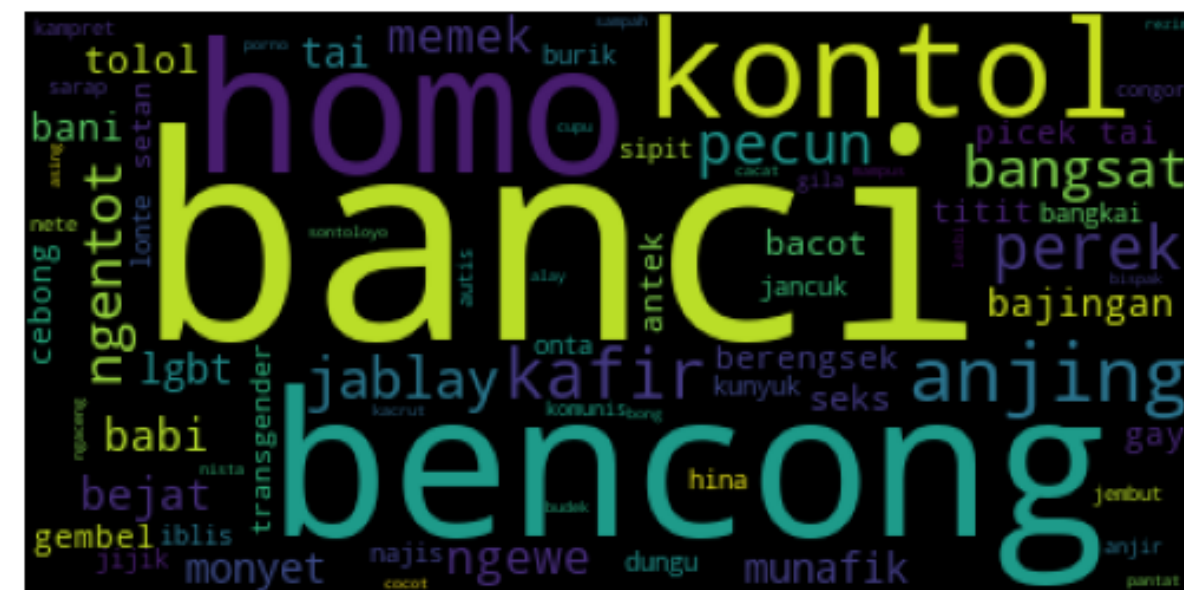
WORDCLOUD

[illegible]

Abusive Word - Race



Abusive Word - Gender



HASIL DAN KESIMPULAN

KESIMPULAN

1

Individual menjadi yang paling banyak dijadikan target *hate speech* namun level *hate speech* nya **Weak**. Sedangkan **Group** menjadi yang paling banyak di level **Strong**.

Topik **Religion** menjadi yang paling banyak dijadikan topik dalam *hate speech* dan juga di level **moderate**. Sedangkan **Race** menjadi yang paling banyak di level **Strong**.

2

3

Tidak terdapat kolerasi antara jumlah kata *abusive* terhadap jumlah kata yang terdapat pada *hate speech tweet*. Semakin banyak jumlah katanya belum tentu semakin banyak mengandung kata *abusive*.

Diantara semua kata **abusive** pada seluruh data tweet, kata '**Asing**', '**Komunis**', '**Cebong**', dan '**Kafir**' menjadi kata **abusive** yang paling banyak dipakai. Terbanyak pada topik **Religion** adalah kata '**kafir**', '**komunis**', dan '**rezim**'. Terbanyak pada topik **Race** adalah kata '**komunis**', '**antek**', dan '**asing**'. Terbanyak pada topik **Physical** adalah kata '**asing**', '**komunis**', dan '**cebong**'. Terbanyak pada topik **Gender** adalah kata '**banci**', '**bencong**', dan '**homo**'.

4



Thankyou

